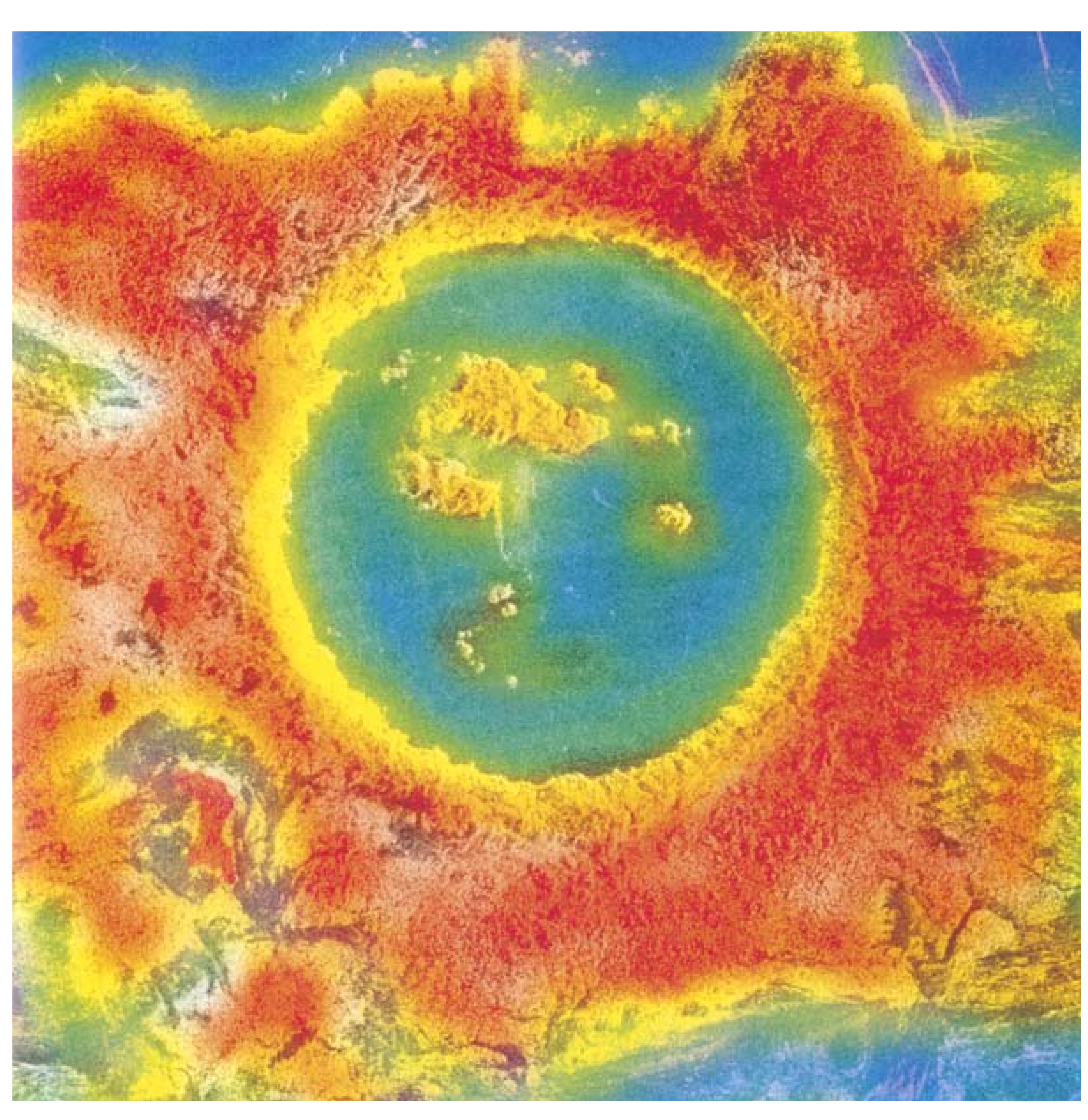
INVESTIGACION y CIENCIA

¿ESTA EN AFRICA NUESTRO ORIGEN?

LA FEMTOQUIMICA

VIVIR DESPUES DE CONGELARSE

Edición española de SCIENTIFIC AMERICAN



VENUS DE CERCA

FEBRERO 1991
600 PTAS.



Febrero de 1991 Número 173

Edición española de **SCIENTIFIC AMERICAN**

SECCIONES

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

J. Vilardell: *Hace...* y *Guerra nuclear fortuita*; Manuel Puigcerver: *Superficie de Venus*; Angel González Ureña: *La formación de las moléculas*; José M. García de la Mora: *El legado de la psicología de la forma*; Joandomènec Ros: *Congelados y vivos*; Antonio Rosas: *¿Está en Africa nuestro origen?*; Santiago Torres: *El cuerpo contra sí mismo*; Luis Bou: *Juegos matemáticos*.

Ciencia y sociedad:

Josep-Enric Llebot y Joandoménec Ros

Ciencia y empresa:

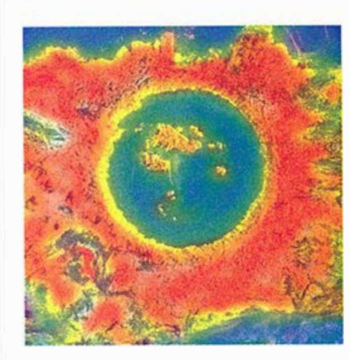
Manuel Puigcerver

Libros:

José Luis Castillo y F. Javier de la Rubia, Josep M. Camarasa, Javier Echevarría, Carlos Gracia y Luis Alonso

PROCEDENCIA DE LAS ILUSTRACIONES

Página	Fuente
6-11	Joe Lertola
12-13	Administración Nacional de Aeronáutica y del Espacio/Laboratorio de Propulsión a Chorro
14	Paul Cochnauer
14-15	Peter G. Ford, Instituto de Tecnología de Massachusetts (*a*), NASA/Laboratorio de Propulsión a Chorro (*b* y *c*)
16-17	NASA/Laboratorio de Propulsión a Chorro
19-26	José Mª Martín Senovilla y Magda Mària
42-43	Eadweard Muybridge, Museo de Arte de Philadelphia (*arriba*), Hank Iken (*abajo*)
44	Instituto de Tecnología de California (*arriba*), Hank Iken (*abajo*)
45-47	Hank Iken
51	Bev Doolittle, © 1979 The Greenwich Workshop, Inc.
52	Frederic Lewis, Archive Films
53-54	Johnny Johnson
55-56	Johnny Johnson
58-62	Patricia J. Wynne
63	Kenneth B. Storey y Janet M. Storey
67	Léopold Marboeuf, Biblioteca Pública de Nueva York, Picture Collection
68-71	Ian Worpole, Christopher B. Stringer/Museo del Hombre, Instituto de Paleontología Humana, Paris
72	Ian Worpole
74	Tomo Narashima
76	CNRI/Science Photo Library, Photo Researchers, Inc. (*arriba*), Joe Lertola (*abajo*)
77	Paul Drake
78-79	Joe Lertola
80	Edward Bell, *New Haven RegisterLertola*
80	Edward Bell, *New Haven Register* (*inset*)
81-82	Joe Lertola
83	Gilbert Faure/SPL, Photo Researchers, Inc. (*izquierda*), Moredun Animal Health Ltd./SPL, Photo Researchers, Inc. (*derecha*)
84	Patricia J. Wynne
85-87	Edward Bell



LA IMAGEN de radar de la portada muestra un cráter de impacto de 50 kilómetros de diámetro en Venus, vista enviada por el vehículo espacial *Magellan* (véase "Superficie de Venus", por R. Stephen Saunders, en este número). El falso color realza los detalles: el azul indica las zonas más lisas y menos reflectoras, mientras que el rojo denota las más abruptas y reflectoras al radar. El cráter parece nuevo y bien definido, lo que prueba que en ese planeta se ha producido escasa erosión pese a la densidad de su atmósfera.

INVESTIGACION Y CIENCIA

DIRECTOR GENERAL Francisco Gracia Guillén
DIRECTOR EDITORIAL José María Valderas Gallardo
DIRECTORA DE ADMINISTRACIÓN Pilar Bronchal Garfella
PRODUCCIÓN César Redondo Zayas
SECRETARÍA Purificación Mayoral Martínez
EDITA Prensa Científica, S. A. Viladomat, 291 6º 1ª - 08029 Barcelona (ESPAÑA)
Teléfonos 321 81 91 - 321 83 48 Telefax 419 47 82

SCIENTIFIC AMERICAN

EDITOR Jonathan Piel
BOARD OF EDITORS Alan Hall, *Executive Editor*; Michelle Press, *Managing Editor*; Timothy M. Beardsley; Elizabeth Corcoran; Deborah Erickson; Marguerite Holloway; John Horgan; Philip Morrison, *Book Editor*; Corey S. Powell; John Rennie; Philip E. Ross; Ricki L. Rusting; Russell Ruthen, Gary Stix; Paul Wallich; Philip M. Yam.
PUBLISHER John J. Moeling, Jr.
ADVERTISING DIRECTOR Robert F. Gregory
PRESIDENT AND CHIEF EXECUTIVE OFFICER Claus-Gerhard Firchow
CHAIRMAN OF THE BOARD Georg-Dieter von Holtzbrinck
CHAIRMAN EMERITUS Gerard Piel

SUSCRIPCIONES

Prensa Científica S. A.
Viladomat, 291 6º 1ª
08029 Barcelona (España)
Teléfonos 321 81 91 - 321 83 48

Precios de suscripción, en pesetas:

	Un año	Dos años
España	6600	12.000
Extranjero	7300	13.400

Ejemplares sueltos:

Ordinario: 600 pesetas
Extraordinario: 775 pesetas

- Todos los precios indicados incluyen el IVA, cuando es aplicable.
- En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.
- El precio de los ejemplares atrasados es el mismo que el de los actuales.

DISTRIBUCION

para España:

MIDESA
Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel. 652 42 00

para los restantes países:

Prensa Científica, S. A.
Viladomat, 291 6º 1ª - 08029 Barcelona
Teléfonos 321 81 91 - 321 83 48

PUBLICIDAD

Madrid: Gustavo Martínez Ovín
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Tel. 409 70 45 - Fax 409 70 46

Cataluña: Marcel Klein
M. K. Publicidad
Ortigosa, 14-16, 3º, D. 20
08003 Barcelona
Tel. 268 45 05 - Fax 268 16 07

OJD · DIFUSIÓN · PUBLICIDAD

ISSN 0210-136X Dep. legal: B. 38.999-76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona
Fotocromos reproducidos por Tecfa, S.A. Almogàvers, 189 - 08018 Barcelona
Imprime Rotographik, S.A. Ctra. de Caldes, km 3,7 - Santa Perpètua de la Mogoda (Barcelona)

Printed in Spain - Impreso en España

Hace...

... cincuenta años

Scientific American: "En distintos puntos del país vienen funcionando comercialmente, desde el primero de enero de 1941, emisoras de radio de un tipo nuevo, cuyas señales sólo pueden captarse con unos receptores especiales. Estas emisoras trabajan con el sistema de frecuencia modulada (FM) que se describía en las páginas de nuestro número de mayo de 1939. Ciñéndonos a lo esencial, la radiodifusión en FM se diferencia de la radiodifusión convencional, o radiodifusión en amplitud modulada (AM), en tres aspectos muy importantes. Primero, y principal quizá, su espectro de frecuencias audibles es más ancho. Segundo, está casi enteramente libre de parásitos, naturales o artificiales. Tercero, no hay interferencias entre las emisoras que trabajan en el mismo canal. Esto no sólo es importante para los radioyentes, a quienes ya no molestarán los silbidos heterodinos cuando empleen receptores en FM, sino que tiene unas consecuencias económicas trascendentales."

"En los aledaños de East Springfield (Ohio) hay un nuevo laboratorio de acústica donde el sonido de la propia respiración se parece al sonido de un pequeño fuelle y el castañeteo de nuestros dedos suena como un tiro de pistola. Se trata de una nueva instalación de Westinghouse que se emplea como laboratorio de verificaciones sonoras para los frigoríficos domésticos. Es la única cámara insonorizada del mundo que no está aislada del exterior cuando funciona. Su construcción recuerda a un laberinto, tal es su tramado de pasadizos con recodos que giran a 90 y 180 grados. En el centro de esa red de pasadizos, fuertemente acolchados, se encuentra la cámara insonora propiamente dicha, un recinto flotante de 9,5 Tm que veinte resortes de acero mantienen en el aire. Los pasadizos acodados actúan de pantallas que destruyen las ondas acústicas reflejadas."

... cien años

Scientific American: "Se anuncia ahora que los experimentos del profesor Hertz acerca de las ondas eléctricas demuestran: (1) que el medio que interviene en la actividad eléctrica es el mismo que sirve de asiento a los fenómenos luminosos; (2) que ambas perturbaciones se propagan bajo las mismas condiciones y a la misma velocidad; (3) que ciertos fenómenos eléctricos y luminosos son de idéntica naturaleza. Ha compuesto además un dispositivo capaz de disminuir la longitud de las ondas luminosas, adelantando así la posibilidad de producir luz eléctrica por procedimientos industriales directos. La lámpara de arco eléctrico desperdicia nueve décimas partes de la energía en rayos caloríficos invisibles. Bastaría con acelerar esos rayos para que aparecieran en forma luminosa; el profesor Hertz espera poder hacer algo al respecto manejando corrientes alternas."

"En las estaciones ferroviarias y en las de transbordadores pueden encontrarse unas máquinas que, por un centavo, lanzan una gota o dos de un líquido que pasa por perfume y que, en muchos casos, sirve de tenue enmascaramiento de la falta de higiene corporal. Este ingenioso artefacto fue patentado el 19 de noviembre de 1889 por Lewis C. Noble de Boston (Massachusetts). Pero, volviendo hacia atrás las páginas de la historia, descubrimos que en el Egipto Antiguo los sacerdotes tenían una fuente de ingresos en la distribución de agua lustral, para lo cual se servían de una vasija de agua purificadora sustancialmente igual a las modernas máquinas expendedoras. Su mecanismo, descrito por Herón hace más de dos mil años, corrobora el viejo aforismo de que "nada hay nuevo bajo el Sol"."

"Si una caída vertical rápida alcanza una magnitud excepcional, producirá unos trastornos psicológicos de la misma clase que los que experimentan las personas en los columpios campestres, toboganes, tiovivos, la visión de abismos, etc. He aquí el terreno a explorar. La instalación está constituida por una torre de varios centenares de metros de altura y un camarín cerrado. En éste penetran los pasajeros y luego se le deja caer libremente desde lo alto de la torre. Al cabo de 100 metros de caída, la velocidad del camarín es de 45 metros por segundo, y de 77 metros por segundo al cabo de 300 metros. Para que esta maniobra sea viable, basta con que los pasajeros lleguen indemnes al final del recorrido. Ello puede conseguirse confiriendo al vehículo la forma de una granada de cañón larga y muy ahusada y recibiéndola en un pozo de agua de profundidad suficiente."

Los espacios en gris
corresponden a publicidad
en la edición impresa

Guerra nuclear fortuita

La modificación de los arsenales de las superpotencias y la de los procedimientos de manejo podría reducir el riesgo de un apocalipsis atómico desencadenado por azar

Bruce G. Blair y Henry W. Kendall

Si en los próximos diez o veinte años estallara una guerra nuclear, sería probablemente por puro accidente. La amenaza de que alguien lance el primero un ataque calculado y a sangre fría parece desaparecer del horizonte, pero bajo la superficie en calma de la diplomacia constructiva entre quienes fueron los antagonistas nucleares tradicionales acecha el peligro de un empleo impremeditado del armamento nuclear. Una utilización fortuita, no autorizada o involuntaria de estas armas se ha convertido en la senda más plausible hacia una guerra nuclear.

Las dos superpotencias, así como Francia, Gran Bretaña y China (miembros añejos del club nuclear), constituyen gérmenes potenciales de un disparo accidental de misiles. La aparición de nuevas potencias nucleares —pensemos en la India, Pakistán e Israel, algunas equipadas con misiles balísticos— confiere carácter de urgencia a la cuestión de las salvaguardas nucleares dentro de los asuntos prioritarios de la seguridad internacional.

La estabilidad política de alguno de los nuevos países con capacidad nuclear es dudosa, por lo que las salvaguardas, físicas o de procedimiento, con que puedan dotar a sus armas podrían ser fácilmente anuladas. No es probable, empero, que un ataque por parte de uno de esos países, sea contra una superpotencia o contra uno de sus aliados, desencadenase una respuesta nuclear generalizada. La máxima amenaza de guerra nuclear impremeditada se encuentra todavía en los arsenales norteamericano y soviético, por lo que la actuación más prometedora encaminada a mitigar los riesgos radica en introducir cambios en esos arsenales.

La probabilidad de una guerra nuclear fortuita se reduciría no poco si en todas las armas nucleares se instalaran seguros codificados a prueba de manipulaciones y si se introdujeran dispositivos que permitieran desactivar los artefactos nucleares, incluso tras su lanzamiento. Además, norteamericanos y soviéticos deberían fiar menos su política en la peligrosa estrategia de respuesta precoz de contragolpe, así como dilatar el tiempo de reacción, o disponibilidad, de sus fuerzas nucleares.

La agitación social y política dentro de la Unión Soviética suscita temores a una guerra nuclear imprevista. La confusión civil acrecienta la posibilidad de que grupos étnicos rebeldes o facciones disidentes se hagan con armas nucleares. Y otra cuestión: la sociedad soviética sufre dislocaciones profundas que podrían llegar a cuartear los cimientos de su sistema de mando nuclear. Aunque EE.UU. no padezca tales convulsiones sociales, su sistema de mando nuclear no excluye cierto riesgo de que el armamento se empleara contraviniendo las intenciones de las autoridades legítimas.

La propia organización de las fuerzas de misiles, tanto en EE.UU. como en la URSS, hace que sea posible lanzar un gran número de misiles sin autorización con la misma facilidad que uno solo. Un fallo de control en el vértice de la cadena de mando o en niveles más bajos (acaso resultante de una ruptura violenta de las instituciones políticas soviéticas) podría acarrear un ataque que causase enormes destrucciones y provocar, quizás, un intercambio nuclear de mayores proporciones.

Si ocurriera un ataque nuclear procedente de la Unión Soviética, en el lanzamiento mínimo podría intervenir un batallón (de 6 a 10 misiles) e incluso un regimiento (de 18 a 30 misiles). Cada misil transporta hasta 10 ojivas y la salva podría componerse de hasta 300 explosiones nucleares. En EE.UU., en el lanzamiento mínimo podría intervenir bien una escuadrilla de 10 misiles, bien un escuadrón de 50, llevando hasta 500 ojivas, cada una de las cuales de una potencia que multiplica por 25 la de la bomba de Hiroshima. Aunque no siguieran represalias, la destrucción y pérdida de vidas consiguientes empequeñecería a cualquier otra calamidad que hasta ahora haya padecido el género humano.

Los comandantes nucleares, tanto americanos como soviéticos, se enfrentan a un inevitable dilema: deben ejercer un control negativo sobre las armas para evitar que se usen mal, pero también deben ejercer un control positivo para asegurar que se empleen cuando se autorice. Ahora bien, las medidas que reducen la posibilidad de lanzamientos indeseados pueden aumentar la posibilidad de que no se cumplan las órdenes legítimas de lanzamiento. Amparándose en ello, los comandantes nucleares se han opuesto a los dispositivos de seguridad, o salvaguardas, más perfeccionados, arguyendo que tales dispositivos debilitarían la disuasión nuclear. Se han tolerado deficiencias en el control negativo y, aunque se han introducido gradualmente algunas

BRUCE G. BLAIR y HENRY W. KENDALL son, respectivamente, investigador en política exterior del Instituto Brookings y profesor de física en el Instituto de Tecnología de Massachusetts. Blair se dedica al estudio de operaciones nucleares, asignación de blancos, mando y control y salvaguardas. Doctor en investigación operativa por la Universidad de Yale, ha prestado servicio oficial, en la sección de lanzamiento de misiles Minuteman, en el Mando Aéreo Estratégico de EE.UU. Kendall es miembro fundador y actual presidente de la Unión de Científicos Preocupados. Su campo de actividad en la física se centra en las interacciones electromagnética y débil entre partículas elementales. En 1990 compartió el premio Nobel de física por el descubrimiento de los quarks.

medidas correctoras, todavía está por encontrar una solución de compromiso satisfactoria.

Una posición alarmista es injustificable; el objetivo primordial en tiempo de paz de los sistemas de mando nucleares, tanto americanos como soviéticos, es impedir de modo fiable el disparo ilícito o accidental de tan siquiera una sola arma nuclear. La jerarquía nuclear sabe muy bien que probablemente no sobreviviría a las repercusiones políticas de un error importante en el cumplimiento de esa función. Los dos bandos han desarrollado sus armas basándose en unos principios de diseño y unos métodos de utilización refinados para asegurarse un eficaz control negativo sobre decenas de millares de ojivas geográficamente dispersas; hasta la fecha han cumplido su misión.

No obstante, tampoco es justificable una confianza absoluta. Ni el más minucioso estudio de las eventualidades que podrían conducir al desastre agota las posibilidades adversas, tal como atestigua la explosión de la lanzadera espacial *Challenger* y el accidente nuclear de Three Mile Island. Además, los puntos más débiles de las actuales salvaguardas aflorarían seguramente en una crisis, circunstancia en la cual ninguna de las superpotencias posee mucha experiencia en mantener el control negativo.

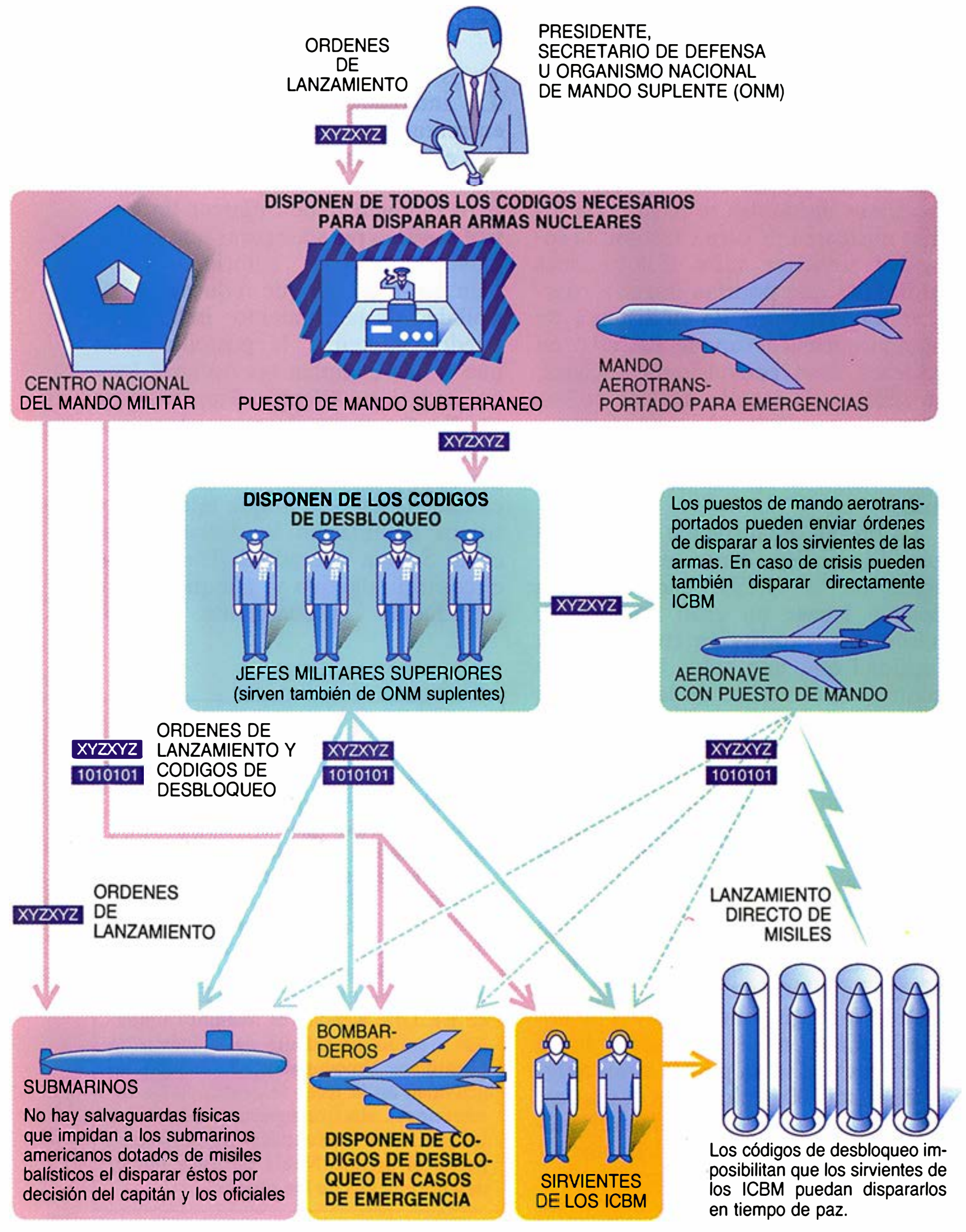
Durante las crisis, la prioridad de los sistemas de mando nucleares se desplaza hacia el control positivo a expensas de las salvaguardas. Cuando en 1968 los soviéticos invadieron Checoslovaquia, pusieron al menos un ejército de sus fuerzas de cohetes estratégicos una raya por debajo de la máxima alerta. De los parques de almacenamiento retiraron las ojivas y éstas las llevaron a los misiles, listos en sus plataformas de lanzamiento. Con estas operaciones se puso en peligro una firme salvaguarda contra el disparo fortuito: la separación física entre las ojivas y sus rampas de lanzamiento.

Durante la guerra árabe-israelí de 1973 los norteamericanos acometieron acciones semejantes. Así, se aumentó el número de bombarderos de largo alcance en estado de alerta en tierra, se hicieron a la mar los submarinos con misiles nucleares a bordo y casi todos los misiles estratégicos terrestres se pusieron a punto para el lanzamiento. Los comandantes de cada arma sacaron de sus cajas fuertes de doble cerradura las llaves de lanzamiento y los códigos presidenciales de lanzamiento, aumentando con todo ello la posibilidad de que se realizaran disparos no autorizados.

Entre los procedimientos vigentes que desde entonces se añadieron a este nivel de alerta se cuenta la activación de enlaces de radiocomunicación especiales que permiten a un avión del mando militar disparar por control remoto la totalidad de los mil misiles balísticos intercontinentales (abreviadamente ICBM, de "Intercontinental Ballistic Missiles") que los americanos tienen en bases terrestres. Esto disminuye la capacidad de las dotaciones de tierra para vetar una orden de lanzamiento ilícita.

Las acciones emprendidas en este tipo de alertas no pueden ser totalmente controladas por los dirigentes políticos. La vulnerabilidad de las fuerzas nucleares y del mismo sistema de mando ante los ataques nucleares engendra cierta presión para delegar en puntos más bajos de la cadena de mando la autoridad sobre las alertas nucleares y el disparo de las armas.

Sin embargo, norteamericanos y soviéticos parecen disentir sustancialmente acerca del grado en que hay que delegar el control positivo durante una crisis. El sistema de mando americano es descentralizado y permite que el comandante de cada arma decrete cualquiera de los procedimientos de alerta previos al disparo. Así, los jefes militares pueden enviar bombarderos a sus apostaderos cercanos al territorio soviético. Y también ordenar el despegue de los puestos de mando aerotransportados, desde los cuales se dictan las órdenes de



lanzamiento en caso de que un ataque destruya las bases asentadas en tierra.

Las órdenes que afectan a la disposición de las fuerzas nucleares discurren por canales estrictamente militares, con una supervisión civil marginal. Además, la documentación histórica deja poco resquicio para la duda: los anteriores presidentes delegaron siempre en jefes militares clave la autoridad para ejecutar planes de guerra nuclear caso de fallar las comunicaciones ante un ataque nuclear comprobado contra EE.UU. Existen pruebas sólidas de que tales disposiciones siguen vigentes. Por ejemplo, son numerosas las instalaciones militares en posesión de los códigos necesarios para autorizar lanzamientos. La fracción del arsenal americano cuyos seguros consisten en cerraduras normales puede ser puesta a punto para el disparo por numerosas instancias dentro de la cadena de mando militar.

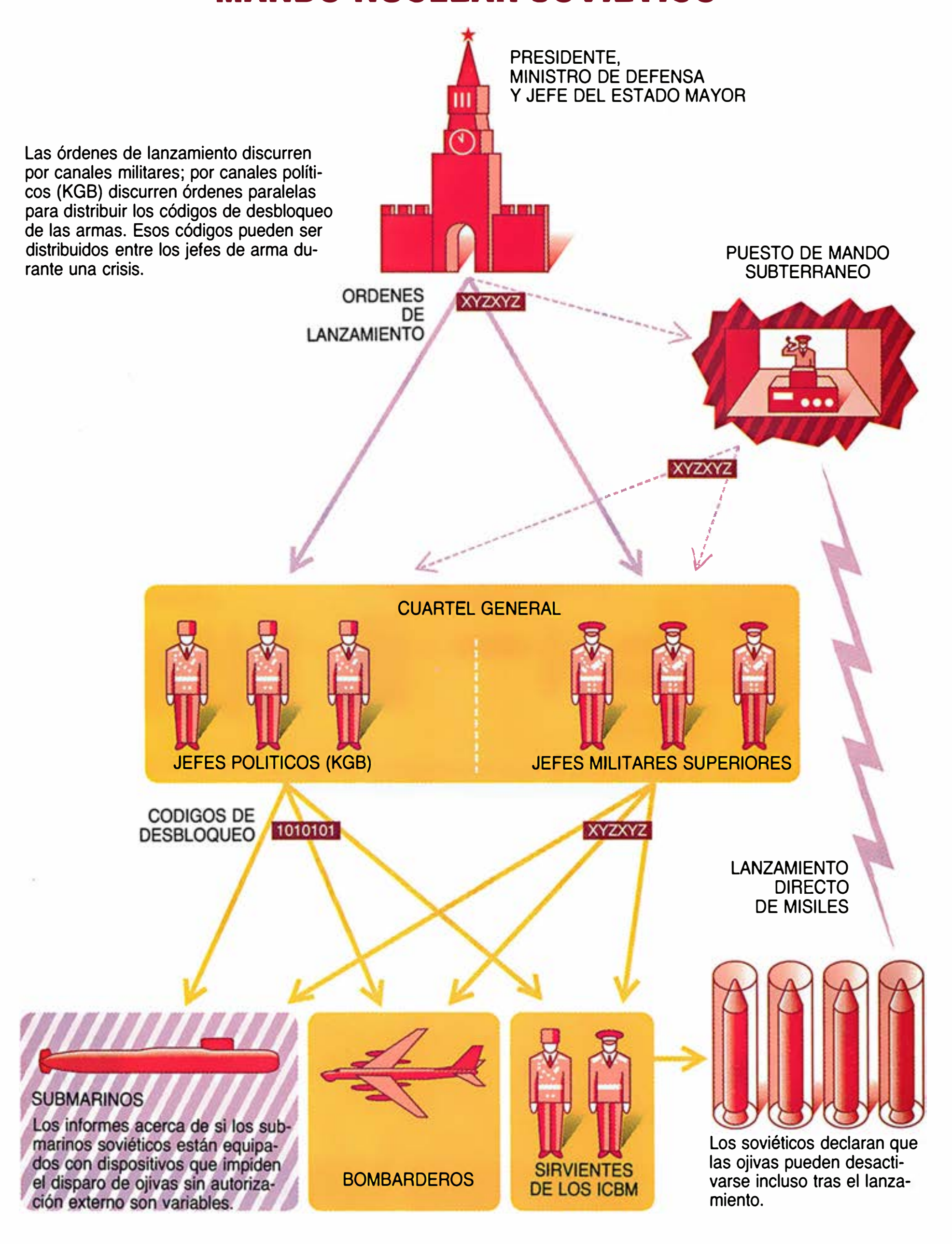
En contraste, ningún jefe militar soviético posee una autoridad de peso suficiente para alertar las fuerzas nucleares, ni mucho menos para ordenar un ataque. Los cambios del estado de alerta requieren la aprobación explícita de los máximos dirigentes políticos. Además, las órdenes respecto a acciones nucleares parece ser que son procesadas en paralelo por varios canales de control independientes, al objeto de asegurar su estricta conformidad con los planes políticos. Hay unidades del KGB (la policía política secreta soviética) que custodian las armas nucleares tácticas y, según se cree, se encargan de difundir los códigos de desbloqueo entre las fuerzas tácticas y la mayoría de las estratégicas, asegurando así la centralización del control político. Sin embargo, si los códigos se distribuyeran como parte de la preparación para la guerra, se ampliaría el campo de posibilidades para un lanzamiento no autorizado.

Un punto débil adicional en la prevención de una guerra indeseada yace en la estrategia de respuesta precoz de contragolpe; impone ésta a los jefes militares que disparen misiles de represalia tras la confirmación de un ataque, pero antes de que detonen las ojivas enemigas. Tanto americanos como soviéticos confían en esa estrategia, que requiere un funcionamiento perfecto de los sensores (de los satélites y los terrestres) y un comportamiento sin errores de las personas.

La respuesta precoz de contragolpe obliga a las autoridades a decidir si hay que disparar, y contra qué blancos, en muy poco tiempo y sin una idea muy clara del ataque que se supone está en marcha. Hay que responder sin un recuento definitivo de las ojivas, sin una idea clara de los blancos amenazados, sin predicciones acerca de los daños y bajas previstas, sin conocimiento de los objetivos del ataque y posiblemente sin manera de saber si el ataque es deliberado, accidental o no autorizado. Y aunque dispusieran de toda esta información, en el poco tiempo que permite la respuesta precoz de contragolpe no les sería fácil a los jefes militares hacerse cargo de la situación y reaccionar con la cabeza fría.

Por citar una situación verosímil, el jefe del Mando de la Defensa Aérea Norteamericana (NORAD, de "North American Air Defense Command") dispondría de sólo tres minutos, desde el instante de la señal inicial de ataque, para juzgar si el continente está bajo fuego enemigo o no. Evidentemente, esta decisión (y las siguientes que deben tomarse durante los 10 minutos que dura el vuelo de los misiles lanzados desde submarinos o los 30 minutos de los ICBM) implica grandes riesgos de un lanzamiento de armas precipitado, a causa de falsas alarmas, errores de cálculo o confusiones.

En EE.UU. es corriente que varias veces al día se dé uno de los llamados eventos de misiles, o sea, señales de un posible ataque. Cuando sucede tal evento, el Director Jefe del NORAD debe entablar una consulta especial con el Mando Aéreo Estraté-



gico y el Pentágono y declarar cuál es su valoración de la amenaza a Norteamérica. Además, cada año millares de señales anómalas de sensores precisan de atención y evaluación urgentes. Entre 1979 y 1984, único período del que se dispone de información oficial, el NORAD evaluó anualmente del orden de 2600 señales de alerta extrañas, de las que una de cada veinte requirieron una valoración ulterior porque parecían entrañar una amenaza.

La mayoría de las falsas alarmas, sean debidas a datos incorrectos, microcircuitos defectuosos u otras averías de funcionamiento, se reconocen enseguida; pero quizás una o dos veces al año una alarma persiste tanto que se decreta una alerta nuclear. El último incidente de este tipo públicamente revelado tuvo lugar en 1980, cuando un microcircuito defectuoso generó señales de un ataque soviético masivo.

En la confusión reinante se declaró una alerta nuclear y el Director Jefe no consiguió emitir a tiempo un informe correcto. (Fue destruido al día siguiente.) La alerta duró más y llegó a un nivel más alto que lo justificable por la situación. En un ambiente de crisis entre las superpotencias, tal confusión podría haber provocado con mayor facilidad que los jefes militares actuaran cuál si realmente hubiérase desatado un ataque.

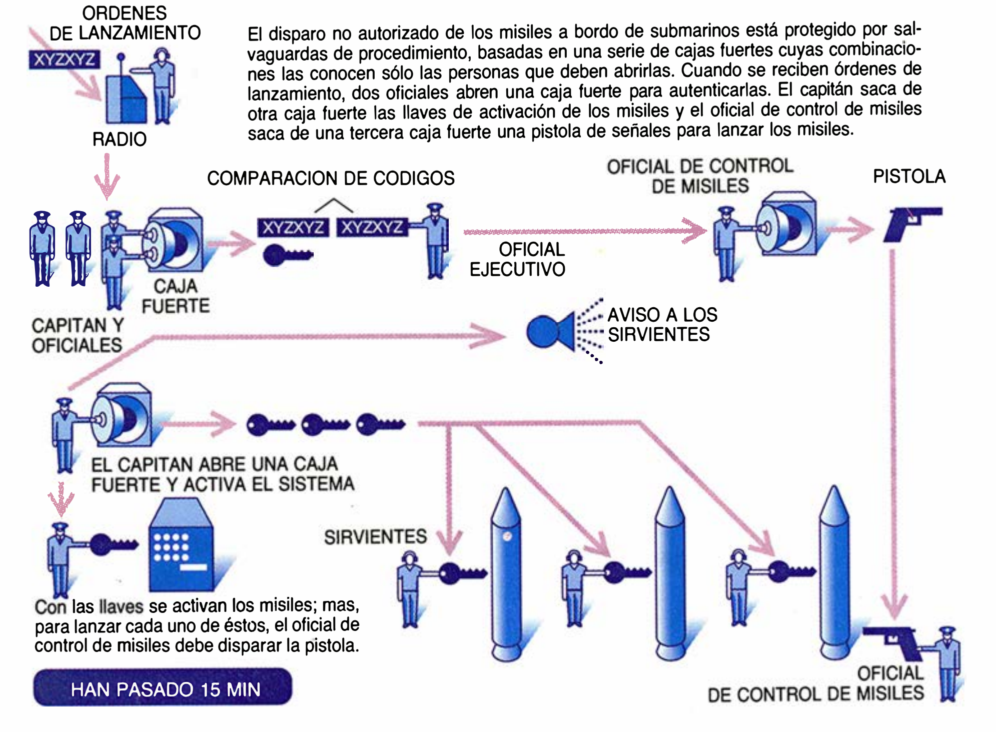
Los procedimientos soviéticos equivalentes para valorar las señales de ataque e iniciar las represalias pecan, parece, de pareja vulnerabilidad a las falsas alarmas. Un general soviético retirado narraba recientemente cómo, en cierta ocasión, fue testigo de señales procedentes de sensores espaciales que alertaban del lanzamiento de misiles americanos Minuteman contra la URSS. Un "operador competente", recordaba el general, determinó que las estelas supuestamente debidas a los escapes de los misiles no eran en realidad sino "motas de luz solar".

Hasta ahora, el personal ha sabido reconocer los fallos de los aparatos que produjeron falsas alertas a tiempo para evitar la guerra. Los mecanismos correctores se han portado aceptablemente, bien que en tiempo de paz cuando los niveles de alerta y ansiedad son bajos.

Aun así, en cada caso fueron desactivadas ciertas salvaguardas contra los lanzamientos indeseados. Por ejemplo, en EE.UU. los sirvientes de lanzamiento de los Minuteman retiraron de las cajas fuertes los códigos de autorización y las llaves de lanzamiento. Las tripulaciones de los bombarderos fueron alertadas y despegaron los aviones de mando sin el conocimiento de las autoridades políticas. Con estas actuaciones se corre el riesgo de provocar reacciones soviéticas, las cuales a su vez podrían reforzar la sensación de ataque inmediato en los americanos. La cuestión de si tales interacciones, basadas en falsas alarmas, podrían desencadenar un ataque nuclear sigue pendiente.

Introduciendo cambios técnicos y de procedimiento se reducirían las posibilidades de guerra involuntaria. En cabeza de la lista figura el uso extensivo de los llamados acopladores facultadores de acción (PALS, de "Permissive Action Links"). Estos seguros, o trinquetes, electromecánicos

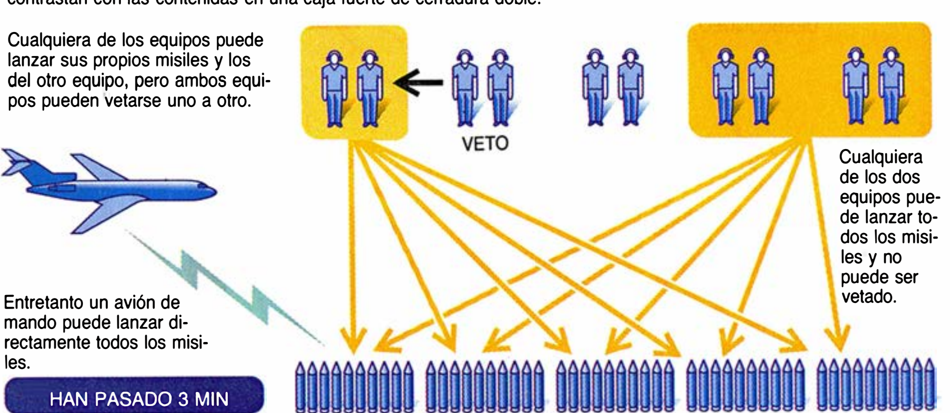
LANZAMIENTO DESDE UN SUBMARINO NORTEAMERICANO



LANZAMIENTO DE UN ICBM NORTEAMERICANO



El disparo no autorizado de los ICBM con base en tierra está protegido por salvaguardas físicas. Los oficiales de lanzamiento reciben las órdenes de lanzamiento y los códigos de habilitación de los misiles. Las órdenes recibidas las contrastan con las contenidas en una caja fuerte de cerradura doble.



impiden que detonen las ojivas a menos que se inserte un código infranqueable. Las autoridades superiores sólo difunden tales códigos entre los jefes de arma cuando el lanzamiento ha sido debidamente decretado. Los PALS impiden los disparos no autorizados a los propios sirvientes del arma y a soldados enemigos o a terroristas que pudieran capturar una ojiva. Hay dispositivos parecidos, llamados sistemas de conmutación codificada, que impiden la apertura de los compartimentos de bombas y el disparo de misiles.

Estos dispositivos fueron adoptados por EE.UU. a comienzos de los años sesenta, instalándolos en las ojivas tácticas asignadas a las fuerzas aliadas acantonadas en ultramar; actualmente, todas las armas tácticas americanas con base en tierra están protegidas por PAL o con sistemas de conmutación codificada. Es de presumir que los misiles soviéticos con base en tierra y sus bombarderos están equipados con dispositivos similares.

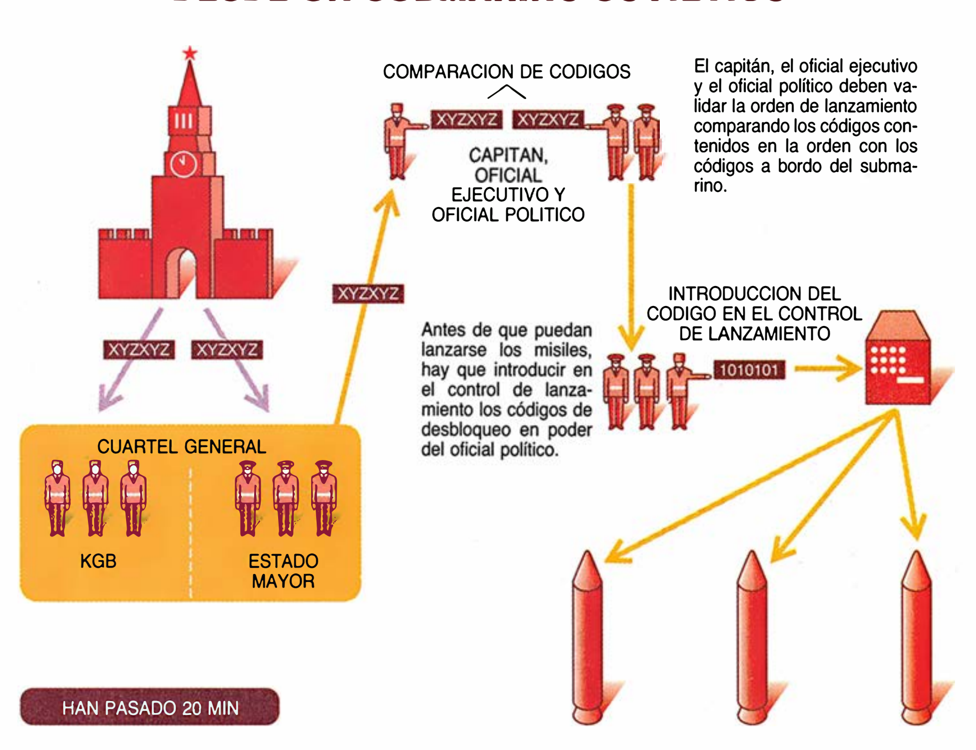
Sin embargo, las fuerzas navales nucleares de ambos bandos no están, por lo general, equipadas con PAL. El consiguiente peligro de lanzamiento accidental reviste especial gravedad en el caso de los misiles crucero lanzados desde el mar. Estas armas son de largo alcance y las transportan buques de superficie y submarinos de ataque que entrarían en combate en las primeras fases de un conflicto. Algunas ojivas británicas y francesas se encuentran en posiciones adelantadas, corriendo el mismo peligro.

Otra manera de reducir el riesgo de guerra nuclear fortuita sería disminuir la rapidez de reacción. Los autores estimamos que entre el 50 y el 80 por ciento de la fuerza soviética de ICBM se mantiene rutinariamente en configuración de "listos para lanzar". En tiempo de paz, el 15 por ciento de su fuerza de misiles balísticos a bordo de submarinos se encuentra desplegada en el mar y ninguno de sus bombarderos de largo alcance se halla en estado de alerta o siquiera cargado con armas nucleares.

Entretanto, EE.UU. mantiene aproximadamente el 90 por ciento de sus ICBM en estado de "listos para lanzar", capaces de hacer fuego en menos de tres minutos. La mitad de la fuerza submarina de misiles balísticos se encuentra navegando en todo momento y la mitad de esos navíos pueden disparar sus misiles en menos de 15 minutos. La cuarta parte de los bombarderos estratégicos se hallan en estado de alerta en tierra listos para entrar en acción en cinco minutos.

Esta elevada rapidez de reacción es un anacronismo en la nueva era de las relaciones EE.UU.-URSS y plantea un peligro innecesario. Los porcentajes de ambos arsenales en estado de alerta deben reducirse a una fracción de los niveles actuales. Sin erosionar el poder disuasorio podría llevarse a cabo una reducción unilateral a la tercera parte, y cabría recortar más en etapas sucesivas bajo un acuerdo verificable.

Además, las ojivas de las unidades que no estuvieran en alerta inmediata deberían ponerse bajo la custodia de organismos civiles, como estaban en los años cincuenta, cuando se asignó a la Comisión de Energía Atómica la custodia de las armas atómicas en tiempo de paz. El papel civil en la ad-





ministración de las reservas de ojivas debería reforzarse para que, en caso de crisis, el armar unidades con capacidad nuclear quedara bajo un estrecho control político. Aunque no faltan analistas que aducen que el riesgo de una interpretación errónea es menor si las fuerzas nucleares mantienen continuamente una elevada capacidad de reacción, esta postura contradice las tendencias actuales en las relaciones entre las superpotencias.

La adopción de niveles de alerta más bajos permitiría retirar las ojivas u otros componentes vitales cualesquiera de alguna porción del arsenal estratégico, evitándose totalmente de ese modo que un lanzamiento accidental produjera una explosión nuclear. Este peligro no es en absoluto ilusorio. Los soviéticos revelaron hace poco que un misil balístico con cabeza nuclear había sido lanzado accidentalmente durante unas operaciones de mantenimiento rutinario. Afortunadamente cayó a corta distancia de su plataforma de lanzamiento.

La mayoría de los misiles llevan instaladas las ojivas (entre ellos, todos los misiles estratégicos de las potencias nucleares de primera fila). La estrategia de respuesta precoz de contragolpe y los actuales niveles de alerta excluyen cualquier otra configuración. Un acuerdo verificable tendente a separar ojivas y misiles (práctica soviética habitual en los años sesenta) reduciría las posibilidades de un eventual lanzamiento fortuito. Además, este cambio de configuración de las armas sería una forma menos radical, y por ello más aceptable, de control de armamentos.

Acoplar en tiempo de crisis los componentes retirados podría interpretarse como una preparación para el ataque, incrementándose así la posibilidad de guerra fortuita. Como contrapartida, no obstante, la disminución de la posibilidad de lanzamientos accidentales y el alivio de tensiones que supone la existencia de menos misiles listos para hacer fuego inmediatamente pesan más que el peligro de percepción errónea durante una crisis. Además, este último riesgo podrían mitigarlo acuerdos que definan las condiciones en que deben volver a montarse las armas y unos procedimientos de alerta nuclear que garanticen el mantenimiento de un control civil firme.

Cualquiera que sea el nivel de alerta, el factor humano desempeña un papel clave en el riesgo de lanzamientos fortuitos. Todos los integrantes de la cadena de armamento nuclear, desde quienes deciden al máximo nivel hasta los oficiales lanzadores, están sujetos a las fragilidades e inestabilidades humanas. Estas fragilidades se ven agravadas por unas condiciones de trabajo que agudizan el aburrimiento y la sensación de aislamiento. (En los submarinos con misiles, al servicio se le añade la tensión nerviosa que supone adaptarse a un "día" artificial de 18 horas durante los dos meses que dura habitualmente una patrulla.) Tales condiciones pueden a veces acarrear graves trastornos de conducta, entre los que cabe citar el consumo de alcohol, la afición a las drogas y la inestabilidad psíquica y emocional.

Así, en 1989, de los aproximadamente 75.000 miembros del personal militar americano que tenían acceso a armas nucleares y componentes afines, hubo que retirar del servicio a cerca de 2400. De éstos, setecientos treinta abusaban del alcohol o las drogas y los demás padecían trastornos psíquicos o emocionales, eran indisciplinados o incurrían en conductas delictivas. Herbert Abrams, de la Universidad de Stanford, recomienda que se vigile más de cerca el consumo de alcohol y drogas entre los soldados con responsabilidades nucleares y que los médicos que los atienden sean conscientes de las peculiaridades de los deberes que pesan sobre ellos. Además de esto, puede hacerse mucho más para aliviar unas condiciones de trabajo proclives al desequilibrio nervioso. (Todos estos problemas y remedios tienen al menos la misma validez para el aparato nuclear soviético.)

Por su parte, los dirigentes mundiales no están menos sujetos a la presión emocional. Pueden caer en crisis de dependencia del alcohol o de las drogas, sea por prescripción o por automedicación. Estos males afligían a Winston Churchill y a Anthony Eden; Richard M. Nixon estaba tan alterado por el asunto Watergate que no pudo participar en el crucial intercambio de información que, en 1973, tuvo como resultado una alarma nuclear total; posteriormente, altos funcionarios del gobierno tomaron precauciones para evitar que pudiera actuar irracionalmente haciendo uso de su condición de comandante en jefe. Incluso un dirigente tan implacable como Stalin sufría graves tensiones durante los períodos de crisis. Estas flaquezas pueden conllevar cambios de conducta, deterioro de la capacidad de juicio e incluso actos irracionales. Algunos psicólogos sugieren que los dirigentes políticos estén bajo control médico.

Si todas estas medidas no lograsen impedir un lanzamiento fortuito, quedan algunas todavía por tomar para mitigar las consecuencias. Una bastante obvia es reducir el número de ojivas que porta cada misil; otra es desarrollar técnicas para destruir los misiles en vuelo. Una tercera medida sería preparar sistemas detectores de lanzamiento que alerten a ambos bandos de los lanzamientos fortuitos. En el caso de los ICBM bastaría con alguna combinación de sensores ópticos y acústicos; los misiles balísticos lanzados desde submarinos y los misiles crucero requerirían pequeños transmisores que enviaran señales a satélites relé. Un sistema de detección así podría incluir un procedimiento de desactivación al iniciarse la crisis.

Las mejoras del "teléfono rojo" (en servicio desde 1963) reducirían el riesgo de percepciones erróneas acerca de lanzamientos fortuitos, armado de misiles y ojivas y otros actos aparentemente hostiles. El enlace entre EE.UU. y la URSS discurre por un satélite americano y uno soviético empleados primordialmente en comunicaciones civiles, con rutas auxiliares de unión por cable y radio. Pero esta configuración no puede asegurar unas comunicaciones fiables en las condiciones adversas bajo las que puede ser más necesaria. Todas las piezas del sistema son vulnerables al ataque nuclear. Además, las rutas por cable ya han sufrido cortes por descuidos unas cuantas veces. Con satélites antirradiación, exclusivos para este cometido, que funcionen en una banda de frecuencias extremadamente alta, se evitaría que la estática nuclear interrumpiera las señales y sería posible proseguir las conversaciones vía teléfono rojo a través de terminales móviles que podrían acompañar a los centros de mando en sus desplazamientos.

Sin embargo, la actuación decisiva consiste en poner en servicio mecanismos que permitan desactivar los misiles una vez lanzados. Costosos intentos de desarrollar sistemas contra misiles balísticos han desembocado en un callejón sin salida. Las estipulaciones del aún vigente Tratado sobre Limitación de Sistemas contra Misiles Balísticos (conocido como Tratado ABM) limitaría a 100 el despliegue de interceptadores con base en Grand Forks (Dakota del Norte). Este sistema (que costaría unos 10.000 millones de dólares) sólo sería capaz de destruir del orden de 50 vehículos de reentrada y sería ineficaz contra los misiles crucero y los misiles balísticos lanzados desde submarinos. Un sistema más capaz, que rebasaría los límites del Tratado ABM, costaría

mucho más. No existe expectativa realista de que tal proyecto salga adelante.

El único medio plausible para detener un misil tras un lanzamiento fortuito es que el país lanzador lo destruya o dé ocasión al país objetivo para que destruya las ojivas antes del impacto; o sea, un sistema de "orden de destrucción". Lo cierto es que, en 1971, la URSS y EE.UU. firmaron un tratado, aún poco conocido, que estipula lo que ha de hacer cada uno en la eventualidad de un lanzamiento fortuito. Este tratado incluye el acuerdo de que el país "responsable del arma nuclear hará inmediatamente todos los esfuerzos para tomar las medidas *necesarias que tornen inofensiva o destruyan* esa arma sin que cause daño".

Un sistema americano típico podría consistir en una clave codificada generada automáticamente en el momento del lanzamiento, para evitar robos. Tras haberse comprobado el lanzamiento fortuito, la clave sería transmitida a las ojivas, mediante satélites relé específicos, desde un sistema de mando y control especial situado en el Centro Nacional del Mando Militar (el Pentágono). Sería asimismo enviada al país objetivo a través del teléfono rojo, acompañada de datos acerca de las trayectorias de las ojivas, para que pudieran realizarse intentos de destrucción hasta el punto de reentrada. Inmediatamente antes de un lanzamiento autorizado el sistema de destrucción podría ser desactivado por unidades de mando especiales, ajenas al proceso de lanzamiento, al objeto de excluir la remota posibilidad de que el adversario fuera capaz de neutralizar un ataque intencionado.

En las mismas ojivas convendría colocar dispositivos destructores para que quienes hayan de tomar decisiones dispongan del mayor tiempo posible. El mejor punto para la destrucción sería en la mitad de la trayectoria de vuelo, donde los daños serían los mínimos.

Revestiría especial interés, sobre todo en tiempos de crisis, que la destrucción o desactivación pudiera verificarse por ambas partes. Los ICBM podrían emitir una radioseñal codificada o una descarga de tiras de papel metalizado («chaff») visible al radar. Otra posibilidad sería hacer estallar las ojivas con el mínimo rendimiento energético extrayendo el tritio antes de la detonación. Los misiles crucero podrían desactivar sus ojivas, subir luego a elevada altitud, emitir una señal codificada y luego volar hacia el océano o hacia el casquete polar ártico, donde se estrellarían o serían destruidos por una carga explosiva.

Aunque un sistema como éste añadiría peso y complejidad a los vehículos nucleares de reentrada, la incomodidad queda de sobra compensada por la reducción del peligro derivado de un lanzamiento fortuito. Pese a ello, ni los militares americanos ni el Departamento de Defensa han mostrado interés alguno en la adopción de sistemas de órdenes de destrucción. Cosa que resulta sorprendente a la vista de la actitud de la URSS: Viktor Karpov, ministro adjunto de Asuntos Exteriores, nos ha asegurado que los ICBM soviéticos ya llevan sistemas de órdenes de destrucción.

El peligro de explosiones nucleares fortuitas y el riesgo aún peor de guerra nuclear, por reducidos que sean, deben rebajarse aún más. EE.UU. y la URSS deben actuar prontamente para llegar a compromisos aceptables. Cada país goza de amplia libertad de acción para moverse con independencia en este terreno; la mayoría de las mejoras en las salvaguardas pueden llevarse a cabo unilateralmente. Los soviéticos, por ejemplo, han trasladado las armas nucleares desplegadas en zonas de disturbios étnicos a otros parques de almacenamiento en la república rusa. Y proyectan desmantelar las fuerzas de ICBM en Kazajstán, la única región étnica que "hospeda" fuerzas estratégicas soviéticas.

Valdría además la pena que ambos gobiernos intercambiaran opiniones acerca de los riesgos especiales que entrañarían los lanzamientos fortuitos y de las medidas para reducir tales riesgos. Convendría ampliar esas conversaciones para incluir a otros países con arsenales nucleares. Las relaciones nunca han sido más propicias para entablar unas conversaciones fructíferas sobre el tema.

BIBLIOGRAFIA COMPLEMENTARIA

CRISIS STABILITY AND NUCLEAR WAR, dirigido por Kurt Gottfried y Bruce G. Blair. Oxford University Press, 1988.

ABORTING UNAUTHORIZED LAUNCHES OF NUCLEAR-ARMED MISSILES THROUGH POST-LAUNCH DESTRUCTION. Sherman Frankel en *Science and Global Security*, vol. 2, n.º 1, pag. 1-20; noviembre de 1990.

REDUCING THE DANGERS OF ACCIDENTAL AND UNAUTHORIZED NUCLEAR LAUNCH AND TERRORIST ATTACK: ALTERNATIVES TO A BALLISTIC MISSILE DEFENSE SYSTEM. International Foundation, 1990.

THE LOGIC OF ACCIDENTAL NUCLEAR WAR. Bruce G. Blair, Brookings Institution (en prensa).

Superficie de Venus

Las imágenes de alta definición revelan mezcla de accidentes, familiares unos, insólitos otros

R. Stephen Saunders

Por su tamaño, masa y distancia al Sol, Venus parece gemelo de la Tierra. La naturaleza de su superficie ha constituido, sin embargo, un misterio durante mucho tiempo. Las nubes plateadas que lo embellecen ante nuestros ojos lo ocultan también por completo.

Para fortuna de la ciencia, las radioondas de longitudes de onda centimétricas perforan las nubes, permitiendo que los radiotelescopios asentados en el suelo y los instalados a bordo de las sondas espaciales puedan reconstruir una tosca representación del paisaje del planeta. Cuando el vehículo espacial *Magellan* comenzó a cartografiar Venus el pasado 15 de septiembre, éste empezó por fin a revelarnos su verdadera naturaleza. La impresión general es que constituye un mundo dinámico, moldeado por procesos muy similares a los desarrollados en la Tierra, que, no obstante, produjeron resultados llamativamente distintos.

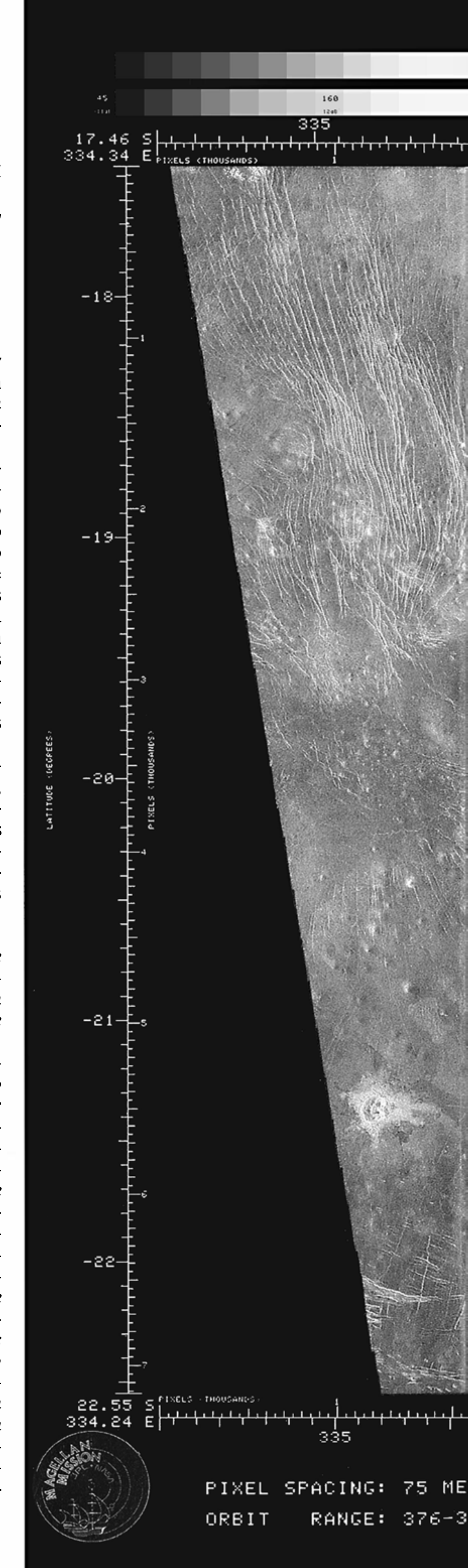
El radar de apertura sintética de la sonda, muy refinado, revela estructuras cuyo diámetro no supera los 120 metros, es decir, la décima parte del tamaño de los que podían detectarse hasta ahora. En las imágenes que presentamos, las regiones accidentadas aparecen brillantes y, oscuras, las regiones llanas, debido a la manera en que las diferentes texturas de la superficie reflejan el radar.

Las dificultades iniciales de comunicación con la sonda *Magellan* se han resuelto en su mayor parte. Aproximadamente el 98 por ciento de las señales llegan ahora sin problemas. Las imágenes de alta definición que ofrecemos configuran una sobria muestra del volumen de información que el vehículo espacial habrá de recoger. *Magellan* cartografiará la totalidad de la superficie de Venus, como ya *Mariner 9* cartografiara Marte en 1971 y 1972. Con notables diferencias. En sólo tres días de operación, *Magellan* envió más datos de los que había remitido la misión *Mariner 9* completa.

Esos tres días de imágenes, que cubren sólo el 1,5 por ciento de la superficie de Venus, se han procesado ya. La riqueza de detalle ha obligado a revisar algunas ideas sobre nuestro vecino. Por lo que se ve, la superficie tiene de 100 a 1000 millones de años de edad; una superficie vieja, pues, a tenor de los criterios terrestres. Sin embargo, esta edad implica que las capas exteriores del planeta se han remodelado y erosionado en época reciente de la historia, de 4500 millones de años, del sistema solar.

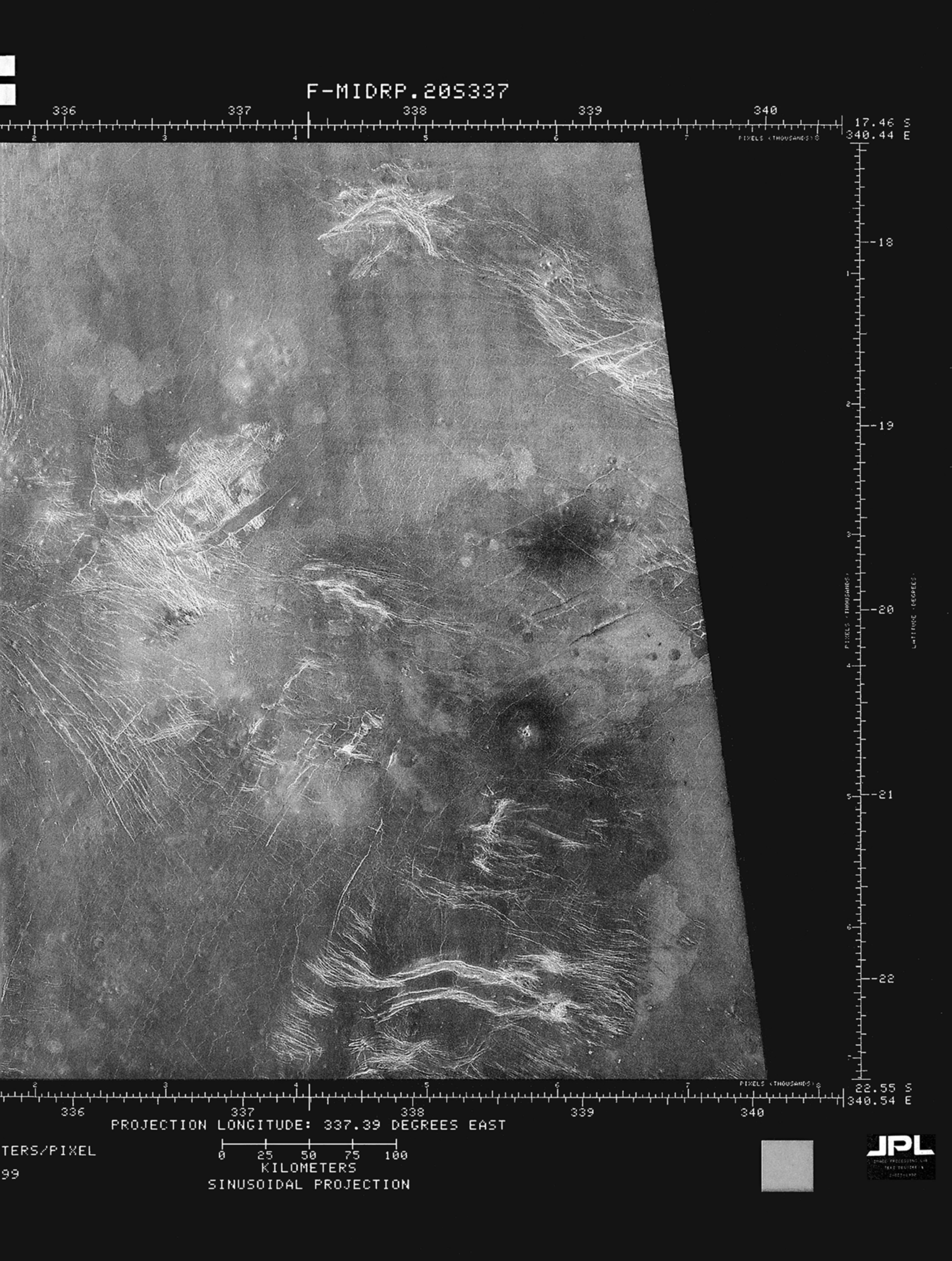
Las calderas volcánicas escalonadas, las extensas corrientes de lava, las cadenas montañosas plegadas sobre sí mismas y las fallas, intrincadas y torturadas, dan testimonio de la actividad interna de Venus. Dudosas siguen siendo, sin embargo, las pruebas

R. STEPHEN SAUNDERS está integrado en el equipo científico del proyecto para la misión *Magellan*, de la NASA, en el Laboratorio de Propulsión a Chorro de Pasadena, California.



1. EN ESTE MOSAICO de radar se aprecia la topografía desigual de Venus. El mapa cubre un área de 475 km de anchura por 545 de longitud, entre las llanuras de Navka y Lavinia.

A lo largo del borde centro-occidental de la imagen hay un agrupamiento de cúpulas volcánicas de diámetros que oscilan entre 1,5 y 7,5 km. Las cúpulas y sus depósitos de lava están situados en la convergencia de las marcas de radar brillantes y lineales, que parecen ser fallas y vaguadas. En algunos puntos, las cúpulas están superpuestas a las fallas. Estas y las vaguadas se extienden hacia el Sur, hasta terminar en llanuras lisas y oscuras; se hallan atravesadas por otras fallas perpendiculares. La configuración nos indica que los episodios de vulcanismo y formación de fallas se han producido muchas veces. Se cree que la región circular y extremadamente oscura en el radar, cerca del borde centro-oriental de la imagen, corresponde a una región de depósitos sedimentarios, relacionados con un impacto meteórico.

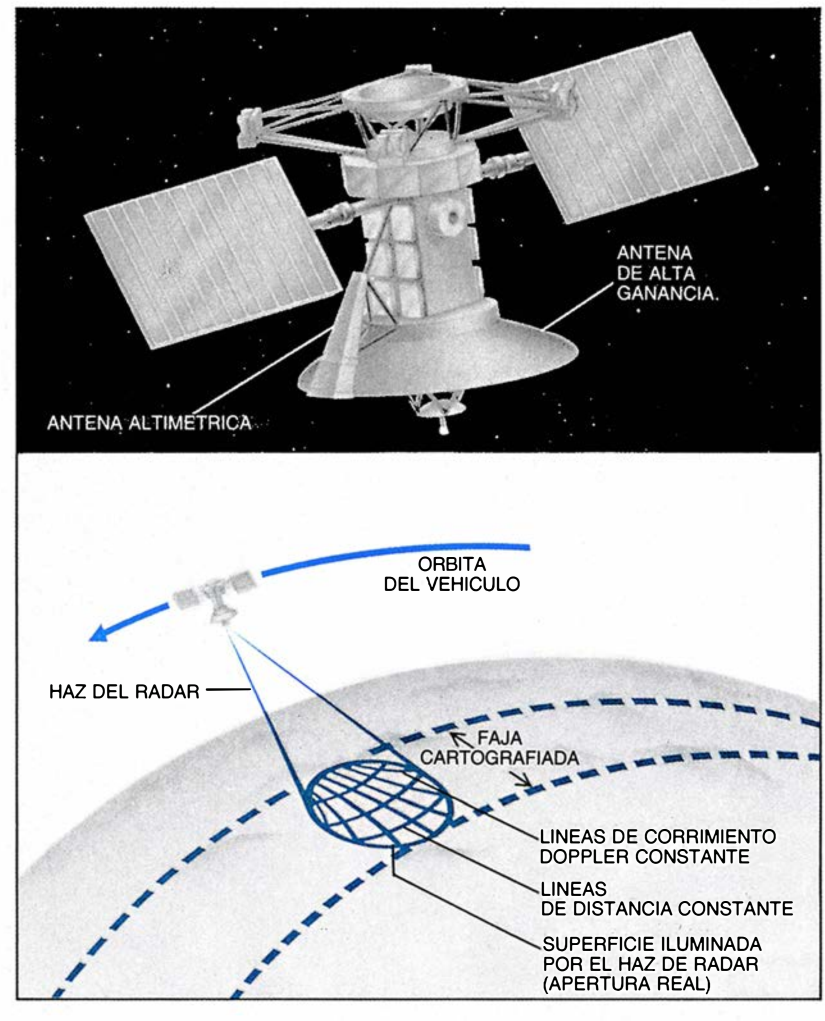


F-MIDRP.20S337
336
337
338
339
340
17.46 S
340.44 E
-18
-19
-20
-21
-22
22.55 S
340.54 E
PROJECTION LONGITUDE: 337.39 DEGREES EAST
TERS/PIXEL
99
0
25
50
75
100
KILOMETERS
SINUSOIDAL PROJECTION
JPL

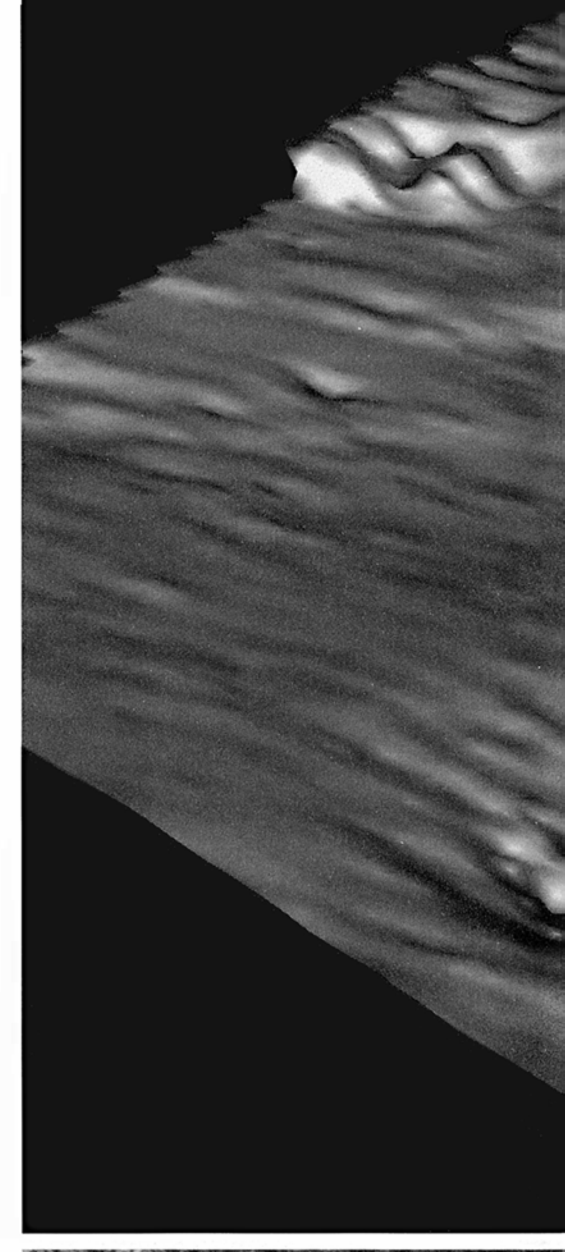
de existencia de placas corticales. Muchos cráteres de impacto parecen rellenos de lava, procedente tal vez de capas ardientes situadas bajo la fina corteza de Venus. A pesar de las muy bajas velocidades del viento en superficie, se distinguen, en varios puntos, señales de erosión y sedimentación arrastrada por el viento.

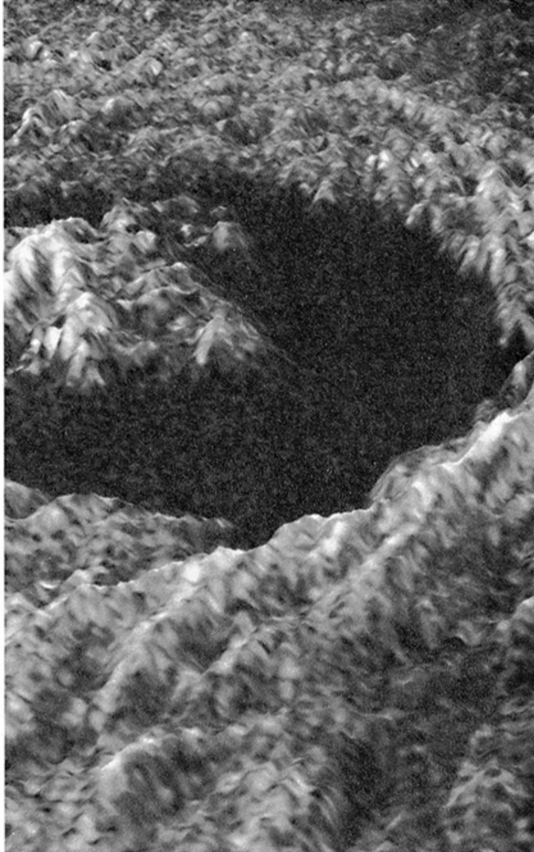
Además de accidentes de aspecto familiar, Venus aloja también algunas extrañas formaciones que a los científicos planetarios nos sumen en la perplejidad. Un número insólito, por lo alto, de cráteres de impacto asimétricos indica que su atmósfera espesa condiciona la llegada a la superficie de los grandes meteoros, así como la distribución de los productos expulsados. Hay accidentes de erosión que recuerdan lechos de ríos, no obstante la imposible presencia de agua líquida a la temperatura de 450 grados Celsius que reina en la superficie. Erosión que sería producida quizá por una lava fluidísima o por gas y polvo procedentes de un impacto. Hay regiones llanas y oscuras, causadas verosímilmente por la explosión de un impacto, que rodean a cierto número de cráteres. Sin embargo, algunas zonas circulares y lisas carecen de cráter central visible.

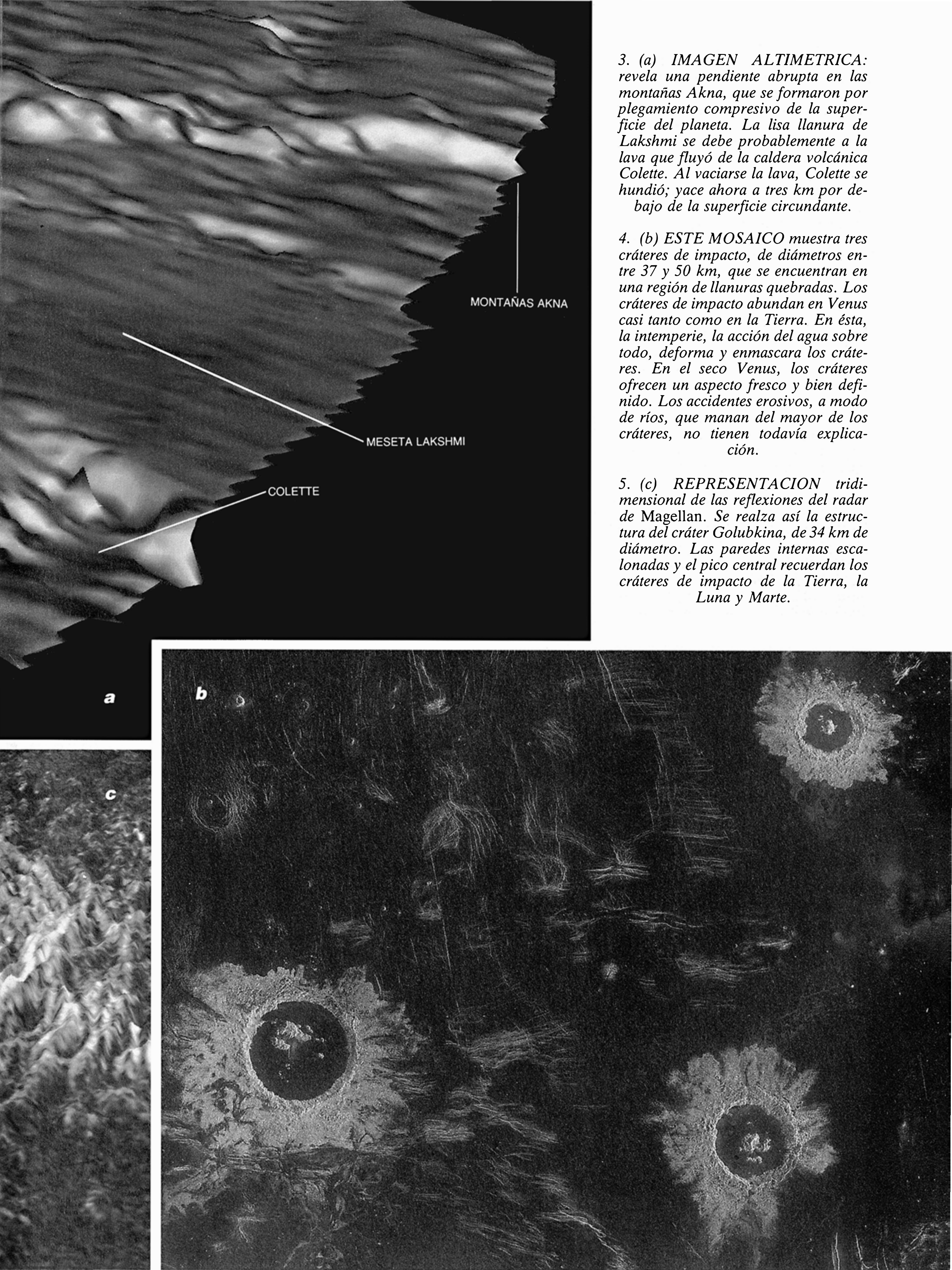
Magellan lleva exploradas sólo las llanuras bajas de Venus. A principios de diciembre, la sonda empezó a cartografiar las tierras altas ecuatoriales. Cabe la posibilidad de que aparezcan accidentes enteramente nuevos. Conforme se vayan examinando fajas cada vez más anchas, podremos interpretar mejor los sistemas de crestas y fallas del planeta. A lo largo de los próximos cinco años, la sonda *Magellan* debe cartografiar ocho veces el planeta, lo que mejorará la resolución y rellenará las zonas perdidas. Permitirá también investigar cambios a corto plazo en la superficie, cambios que podrían mostrar que la inquietud geológica de Venus es, en gran parte, un fenómeno contemporáneo.



2. ANTENA DE ALTA GANANCIA de la sonda Magellan. *Cartografía una faja de Venus durante cada órbita y reenvía los datos a la Tierra. Se combinan múltiples barridos de radar para producir mosaicos. La antena altimétrica mide el tiempo que tardan los impusos del radar en volver al vehículo espacial; de todo esto puede deducirse la topografía de la superficie.*







3. (a) IMAGEN ALTIMETRICA: revela una pendiente abrupta en las montañas Akna, que se formaron por plegamiento compresivo de la superficie del planeta. La lisa llanura de Lakshmi se debe probablemente a la lava que fluyó de la caldera volcánica Colette. Al vaciarse la lava, Colette se hundió; yace ahora a tres km por debajo de la superficie circundante.

4. (b) ESTE MOSAICO muestra tres cráteres de impacto, de diámetros entre 37 y 50 km, que se encuentran en una región de llanuras quebradas. Los cráteres de impacto abundan en Venus casi tanto como en la Tierra. En ésta, la intemperie, la acción del agua sobre todo, deforma y enmascara los cráteres. En el seco Venus, los cráteres ofrecen un aspecto fresco y bien definido. Los accidentes erosivos, a modo de ríos, que manan del mayor de los cráteres, no tienen todavía explicación.

5. (c) REPRESENTACION tridimensional de las reflexiones del radar de Magellan. *Se realza así la estructura del cráter Golubkina, de 34 km de diámetro. Las paredes internas escalonadas y el pico central recuerdan los cráteres de impacto de la Tierra, la Luna y Marte.*

6. (a) CRATER RENIFORME, único por su morfología en todo el Sistema Solar. Quizás un meteoro incidente se rompió en pedazos al atravesar la densa atmósfera de Venus; gruesos fragmentos de materia chocarían casi simultáneamente y dibujarían esa forma irregular.

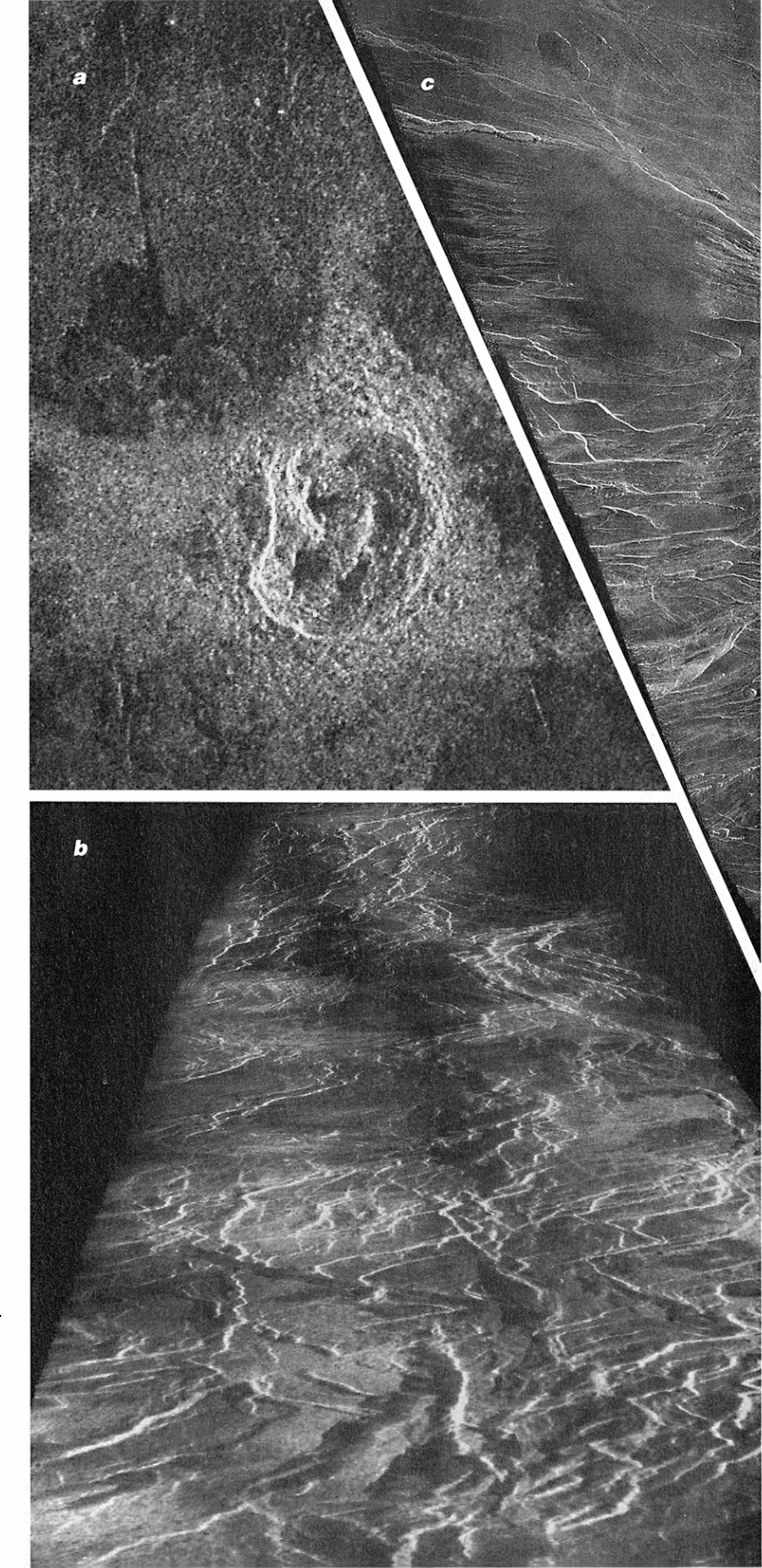
El cráter mide 9 por 12 km; no se encuentran en Venus cráteres de menos de seis km de diámetro, por la presumible razón de que los meteoros pequeños se desintegraron en la atmósfera.

7. (b) LLANURAS BAJAS y jóvenes, desde el punto de vista geológico, abundan en complicados sistemas de cañones, que se formaron cuando la corteza se dilató y abrió por ciertos puntos. Algunos de los cañones se rellenaron después con corrientes de lava. Los cañones poseen anchuras medias de 5 a 10 km y longitudes de 50 a 100 km; su profundidad es de unos 100 metros. Esta perspectiva simula una vista hacia el Norte, con un ángulo de inclinación de unos 10°.

8. (c) FALLAS Y DEPRESIONES deformadas se extienden a lo largo de la mitad inferior de esta imagen de la región de Lavinia. Las estructuras tienen una anchura que oscila de 70 metros a varios kilómetros. Sus extremos septentrionales se ensanchan y forman fosas alargadas. Estos accidentes nacieron quizá cuando el magma subsuperficial chocó con barreras de roca y escapó, hundiendo la corteza suprayacente.

El gran accidente circular parece ser una corona, sede de un afloramiento a gran escala de magma ardiente del interior de Venus. Las líneas de falla que irradian de la corona débense a las tensiones creadas cuando el magma llegó a la superficie. Las coronas son análogas a los puntos calientes volcánicos en la Tierra, tales como el que dio lugar a las islas Hawai. Esta imagen cubre una superficie de 165 km de anchura por 575 de longitud.

9. (d) AL VULCANISMO explosivo se le atribuye el depósito, brillante en el radar, que se extiende a lo largo de 10 km desde el cráter de un kilómetro de anchura situado en el centro. El dibujo rayado de las llanuras circundantes pierde nitidez conforme nos acercamos al cráter, lo que indica que el espesor del depósito es máximo en la proximidad de aquél. La forma del depósito sugiere que los vientos locales transportaron la materia eyectada hacia el Sur, o bien que gradualmente la erosionaron hasta provocar su desaparición salvo en la porción situada en la sombra aerodinámica del volcán.





Singularidades en relatividad general

Existe la creencia generalizada de que las singularidades espacio-temporales –tales como la gran explosión inicial– son consustanciales con la teoría de la relatividad general. Pero no hay razones para sostener dicha afirmación

José María Martín Senovilla

Entre las muchas y diversas contribuciones revolucionarias y geniales de Albert Einstein (1879-1955) a la física se encuentra la teoría de la relatividad general. No se debe confundir a ésta con la relatividad especial. La segunda no es más que una teoría en la que se relacionan diferentes observadores inerciales, mientras que la primera es una teoría de la gravitación o, más precisamente, de la interacción gravitatoria. Conocemos cuatro interacciones fundamentales, a saber, las nucleares débil y fuerte, la electromagnética y la gravitatoria. Entre todas ellas, y con mucho, la gravitatoria es la más débil. Sin embargo, esto no le otorga una categoría inferior; todo lo contrario: se trata de la única relevante para el estudio de fenómenos a gran escala, para fijar ideas a escalas planetarias o mayores.

La causa de la prevalencia de la fuerza de la gravedad es que, por una parte, las interacciones nucleares son de corto alcance –para tamaños del orden del núcleo atómico– y no revisten interés para fenómenos planetarios, estelares o cosmológicos. Por otra parte, aunque la interacción electromagnética es de alcance ilimitado, las partículas que crean el campo electromagnético pueden tener carga positiva o negativa; el efecto que producen las unas es contrarrestado por el causado por las otras, al menos a gran escala, donde habrá, más o menos, el mismo número de cargas positivas y negativas. Queda, por tanto, la interacción gravitatoria; además de ser de alcance ilimitado, se halla ésta generada por las masas (en rigor, por la energía); dado que sólo se conocen masas y energías positivas, la gravedad es siempre atractiva, y en las ocasiones en las que existe una gran concentración de materia se convierte en la interacción cuyos efectos son dominantes.

Como corolario evidente de lo anterior, podemos afirmar que la relatividad general (RG) es la teoría apropiada para el estudio de los fenómenos cosmológicos –evolución del universo–, astrofísicos –estudio de galaxias, estrellas, etc.– y planetarios –órbitas de los planetas en el sistema solar–. Aunque la cosmología y la astrofísica usan también otras ramas de la física, inevitablemente han de tomar como base la RG. Las predicciones de la RG en todos estos campos han sido útiles y, en ocasiones, han recibido una espectacular confirmación; pero lo que ni los más optimistas podían imaginarse era que se pudieran formular y demostrar los denominados teoremas de singularidades. Estos teoremas, desarrollados principalmente por Stephen W. Hawking, de la Universidad de Cambridge, y Roger Penrose, de la Universidad de Oxford, predicen, bajo condiciones determinadas, la existencia de singularidades del espacio-tiempo, o sea, de la base de nuestro mundo clásico.

Dos tipos de singularidades son muy famosos: la gran explosión o "big bang", en la cual se supone que se originó nuestro universo, y la que se produce en el interior de un agujero negro como resultado del colapso gravitatorio de estrellas de gran masa. Hay un tercer tipo de singularidad, menos conocido, de enorme interés: la que se produce cuando chocan dos ondas de gravedad y se destruyen mutuamente de forma catastrófica. Más adelante hablaremos de todas ellas con cierto detalle.

Pero, ¿qué es una singularidad? Esta es una pregunta singularmente difícil de responder. Aun cuando nuestra noción intuitiva es muy clara, plasmar dicha noción en un enunciado riguroso en RG no resulta fácil. Baste por el momento con tales aproximaciones intuitivas. Para ello, nada mejor que recordar que el campo eléctrico creado por una carga aislada es inversamente proporcional al cuadrado de la distancia a la carga. En su expresión matemática, ese campo equivale a uno partido por r al cuadrado ($1/r^2$), donde r es la distancia. A medida que disminuye la distancia, aumenta el valor del campo y, en el límite, cuando la distancia se anula ($r = 0$) el campo se hace inconmensurablemente grande, o sea, infinito. Esta es una singularidad típica, y se pueden encontrar ejemplos similares en todos los campos de la física. En cualquier caso, esta singularidad podría calificarse de benigna, por las siguientes dos razones: primero, es evidente que la singularidad indica la existencia de una fuente del campo –la carga que crea el campo–; segundo, aunque el campo eléctrico se hace infinito, el resto de las magnitudes físicas no se ven afectadas por ello y, aún más importante, las bases fundamentales de toda la física, el espacio y el tiempo, permanecen inalteradas.

La situación en RG es radicalmente diferente, en particular en lo concerniente a la segunda razón citada. Ello se debe a que, en esta teoría, el propio espacio-tiempo guarda íntima relación con el campo gravitatorio. Así, cuando este último se hace infinito, el propio espacio-tiempo deviene singular. Hablando llanamente, el es-

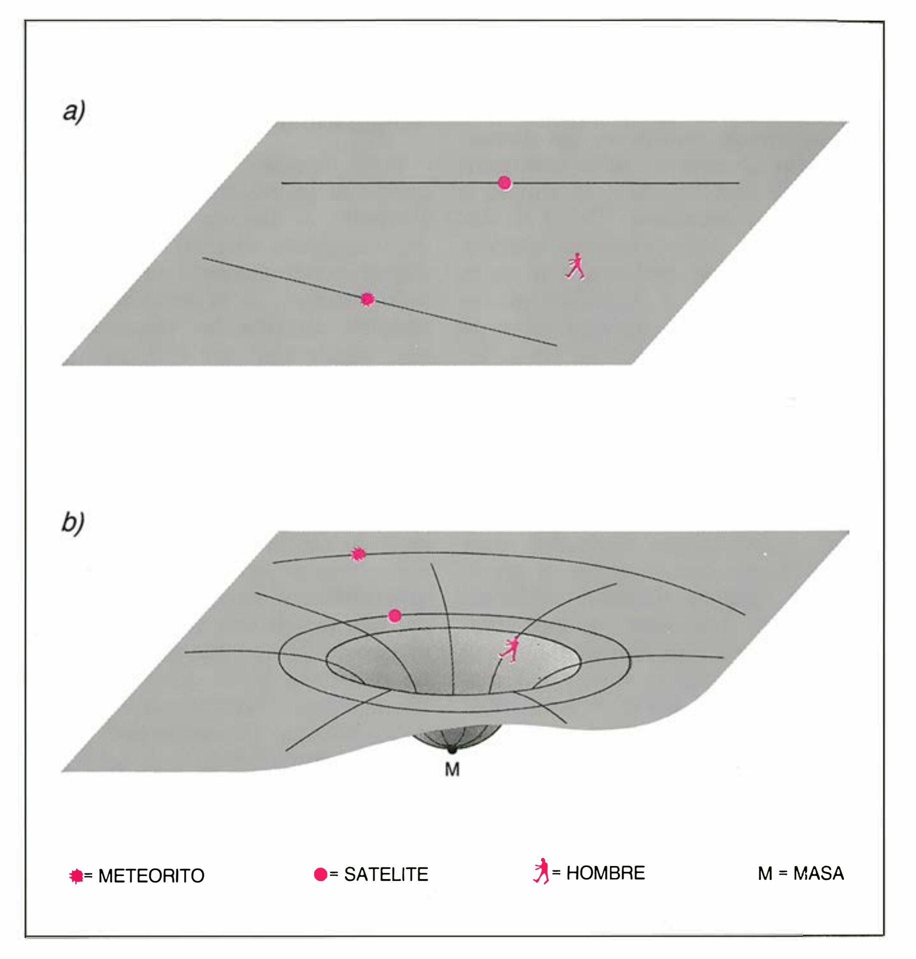
JOSE MARIA MARTIN SENOVILLA se formó en la Universidad de Salamanca, por la que se doctoró en ciencias físicas. Realizó estudios de posgraduado en el Reino Unido. Desde 1990 es profesor de física teórica en la Universidad de Barcelona, donde prosigue su investigación en relatividad general y cosmología teórica.

pacio-tiempo queda destruido y, por consiguiente, todas las leyes de la física pierden su validez. Tamaña catástrofe es, como se comprenderá, inaudita: la propia física predice, a través de la RG, su autodestrucción. Para darnos buena cuenta de ello, nada mejor que un somero repaso de las nociones fundamentales de la RG.

Geometría y causalidad

Cualquier teoría de la gravitación ha de incorporar, entre sus principios esenciales, la propiedad experimental de que todos los cuerpos caen con la misma aceleración en un campo gravitatorio dado. Se atribuye a Galileo el acierto de haber reconocido este importantísimo resultado, según el cual la aceleración que alcanza un objeto grande y pesado –digamos una persona– cuando es arrojado desde cierta altura es la misma que la que alcanza cualquier otro objeto –digamos una pluma de ave–, con independencia de su masa y de sus propiedades internas. (Se han despreciado los efectos de resistencia y rozamiento del aire; el resultado es riguroso en el vacío.)

La profunda comprensión de este hecho condujo a Einstein a enunciar el principio de equivalencia, soporte de su teoría de la RG y que establece la equivalencia local de un campo de gravedad con un sistema de referencia acelerado, no inercial. Un sistema de referencia se describe mediante cantidades geométricas; por tanto, a través de la equivalencia, un campo gravitatorio define la geometría del espacio-tiempo.



1. GEOMETRIA Y MOVIMIENTO EN GRAVITACION. En la figura los cuerpos se mueven sólo por razones geométricas –por la ley del mínimo esfuerzo– y, por tanto, sus trayectorias no dependen de su masa. Si la geometría es plana –en este caso un plano de dos dimensiones– entonces los cuerpos están quietos o trazan líneas rectas. La presencia de una masa *M* genera curvatura del espacio y por tanto la geometría ya no es plana. El movimiento de los cuerpos traza líneas geodésicas, trayectorias circulares o trayectorias abiertas. Ejemplo de las líneas geodésicas sería el de un objeto que cae desde cierta altura en el campo gravitatorio terrestre; un satélite en órbita alrededor de la Tierra dibujaría trayectorias circulares. Ejemplo de trayectorias abiertas pero con una especie de curva sería el de un meteorito que pasa cerca de la Tierra y se desvía por la atracción gravitatoria de ésta.

Es intuitivamente claro que la explicación geométrica del movimiento de los cuerpos satisface la propiedad hallada por Galileo, pues éstos describen trayectorias siguiendo los caminos geométricos más favorables independientemente de su masa. Imaginemos, por ejemplo, que el espacio es una sábana que tenemos extendida y estirada en el aire. Si lanzamos una canica, ésta se moverá por la sábana según una línea recta. Pongamos ahora un cuerpo pesado en el centro de la sábana. Esta se curva –¡su geometría cambia!– y la canica describirá ahora trayectorias curvas. Podemos incluso imaginar un cuerpo tan pesado que rompe la sábana, un hecho ciertamente singular.

De la misma manera, la presencia de energía produce curvatura en el espacio-tiempo, y, por consiguiente, define la estructura geométrica de éste. Una consecuencia directa de todo lo dicho es que la luz también es atraída por la gravedad, y los rayos luminosos siguen trayectorias curvas en un campo gravitatorio. A la luz le gusta ir por los caminos más descansados, y éstos son, por supuesto, los geodésicos definidos por la geometría espacio-temporal. Históricamente, la comprobación experimental de la desviación de los rayos luminosos que pasan cerca del Sol, realizada por Arthur Stanley Eddington y colaboradores en 1919, dio el espaldarazo decisivo para la definitiva aceptación de la RG.

Como se sabe, la velocidad de la luz en el vacío –aproximadamente 300.000 kilómetros por segundo– es la máxima velocidad de propagación de energía. En este sentido, la luz nos proporciona la estructura causal de nuestro mundo. La velocidad de la luz nos ofrece información sobre qué sucesos guardan mutua relación causal. (Se dice que dos sucesos guardan relación causal si uno de ellos es capaz de influir en el otro.) El espacio-tiempo plano de Minkowski –base de la relatividad especial– posee la estructura causal más simple y la considerada, en general, como natural. Hemos visto antes que la presencia de un campo gravitatorio desviaba los rayos de luz; en virtud de ello, configura una estructura causal diferente de la de la geometría plana. En particular, los conos de luz se inclinan y el futuro de cualquier suceso queda modificado. Una representación de estos fenómenos se da en la figura 3, donde se ha exagerado un poco el campo gravitatorio de la Tierra. Estos cambios de la estructura causal son decisivos para el desarrollo de las singularidades, como veremos en seguida.

¿Hasta dónde puede cambiar un campo gravitatorio intenso la estructura causal? Hay múltiples respuestas a esta cuestión. Pero quizá lo más llamativo sea que la respuesta más radical fue ya dada por Pierre Simon

Laplace nada más y nada menos que ¡en 1798! Laplace, mucho antes del advenimiento de la RG, demostró que la fuerza atractiva [gravitatoria] de un cuerpo pesado podría ser tan grande que la luz no sería capaz de fluir fuera de él. Su razonamiento es simple y claro [*véase recuadro*]. De él se deduce que un cuerpo esférico concentrado en un radio menor que su radio de Schwarzschild, r_s, impedirá que las cosas –luz incluida– escapen fuera de tal radio. Para hacernos una idea del orden de magnitud de r_s, señalemos que tiene un valor de 1 centímetro para la Tierra y de 3 kilómetros para el Sol. En otras palabras, si toda la Tierra estuviera concentrada en una esfera de menos de un centímetro de radio, quedaríamos atrapados irreversiblemente en su campo de gravitación y, además, seríamos invisibles desde fuera. Las cosas son más complejas que como se representan en el diagrama de la figura 4: la propia materia del cuerpo siente la irresistible atracción que ejerce sobre sí misma y tiende a caer hacia el centro. Más adelante veremos un diagrama espacio-temporal de la situación real, pero antes debemos adentrarnos en la historia de la RG.

Singularidades y agujeros negros

Poco tiempo después de la publicación de las ecuaciones de la RG por Einstein, el alemán Karl Schwarzschild encontró en 1916 la solución de dichas ecuaciones para un cuerpo aislado, estático, de masa total M y con simetría esférica. Su solución tenía una singularidad en $r = 0$; no le preocupó mucho, pues pensó que era equivalente a la que aparecía en otros campos de la física clásica. Presentábase, sin embargo, otra zona singular en $r = r_s = 2GM/c^2$; de ahí el nombre de radio de Schwarzschild. Después se sabría que la zona definida por $r = r_s$ no constituía ninguna singularidad real, sino un horizonte, esto es, una membrana unidireccional en el espacio-tiempo que permite el paso de partículas hacia dentro, pero no hacia fuera. Preocupado por esta zona extraña, Schwarzschild calculó una solución en el interior del cuerpo que creaba el campo, y concluyó, no sin cierta satisfacción, que tal cuerpo nunca podría rebasar su radio de Schwarzschild.

Las cosas parecían volver a su lógico orden. Posteriormente, los premios Nobel Subrahmanyan Chandrasekhar, de la Universidad de Chicago, y Lev Landau, de la Universidad de Moscú, en 1931 y 1932 respectivamente, establecieron un límite superior para la masa de las estrellas degeneradas –enanas blancas–. El significado y las implicaciones de este resultado son obvios, ya que si una estrella, en el curso de su evolución, llegara al estado de enana blanca con una masa mayor que la dada por el límite de Chandrasekhar, entonces no podría permanecer en equilibrio en tal estado y continuaría contrayéndose. En 1939, Oppenheimer junto con Volkoff y Snyder extendió el trabajo de Chandrasekhar a estrellas de neutrones; si una estrella llega a su estado final de evolución con una masa mayor que cierto límite, entonces seguirá contrayéndose indefinidamente, concluyó. Sus cálculos indicaban que ciertas estrellas dotadas de gran masa podían colapsar y sobrepasar su radio de Schwarzschild. Todo ello era una inevitable consecuencia de que la gravedad es siempre atractiva.

El camino hacia los agujeros negros quedaba así nuevamente abierto. Una vez aceptado que ciertas estrellas pueden colapsar más allá de su radio r_s, sólo faltaba comprender qué pasaba en aquella zona extraña que aparecía en la solución de Schwarzschild para $r = r_s$. Después del trabajo pionero de Eddington, fue en los años 1958-60 cuando se entendió por fin la estructura global de la citada solución, a través de los artículos publicados por Finkelstein, Fronsdal, Kruskal y Szekeres. Todos ellos pusieron de manifiesto la naturaleza de horizonte de la región $r = r_s$, y de qué forma esta región separaba dos zonas diferenciadas, una interna que se denomina agujero negro y una externa más o menos normal. Cualquier observador en la zona externa puede entrar en el agujero negro, pero ningún observador ni señal alguna pueden salir del agujero a la zona externa.

Así pues, los colapsos gravitatorios predichos por Chandrasekhar y Oppenheimer conducían a la aparición de los agujeros negros. Pero, además, de todos estos estudios se desprendía que cualquier tipo de materia, incluida la estelar, caería irremediablemente al centro $r = 0$, una vez se encontrara en el agujero negro. Se producía así una región ($r = 0$) donde la densidad de materia era infinita y el espacio-tiempo se rompía y desaparecía: una singularidad. Parecía, por tanto, que el primer tipo esencial de

El razonamiento de Laplace sobre la relación causal

Desde un punto de vista moderno, el cálculo de Laplace no es sino un cálculo de balance energético. Sea un cuerpo esférico pesado de masa M y radio R. La energía potencial de una partícula pequeña de masa m en la superficie del cuerpo está data por

$$E_p = GMm/R \qquad (1)$$

donde G es la constante de la gravitación universal de Newton. Si la partícula desea escapar del campo gravitatorio del cuerpo pesado, necesitará una energía cinética cuando menos igual a su energía potencial. Pero la energía cinética de la partícula es

$$E_c = mv^2/2 \qquad (2)$$

donde v es su velocidad. Así pues, igualando (1) y (2) obtenemos el valor mínimo de la velocidad que es necesario para que la partícula escape:

$$v_e^2 = 2GM/R \qquad (3)$$

Este valor se denomina velocidad de escape. Como vemos, la velocidad de escape depende sólo de la masa M y el radio R del cuerpo pesado. No depende en absoluto de la masa m de la partícula, lo cual no es más que una consecuencia del principio de equivalencia. Ahora bien, de (3) se sigue que, si la masa del cuerpo aumenta o su radio disminuye, u ocurren ambas cosas, la velocidad de escape aumenta. Pero también sabemos que la velocidad máxima de propagación es la de la luz. Podemos, consiguientemente, preguntarnos cuál debería ser el radio del cuerpo de tal forma que nada, ni siquiera la luz, pudiera escapar. De (3) es inmediato ver que dicho radio es

$$r_s = 2GM/c^2 \qquad (4)$$

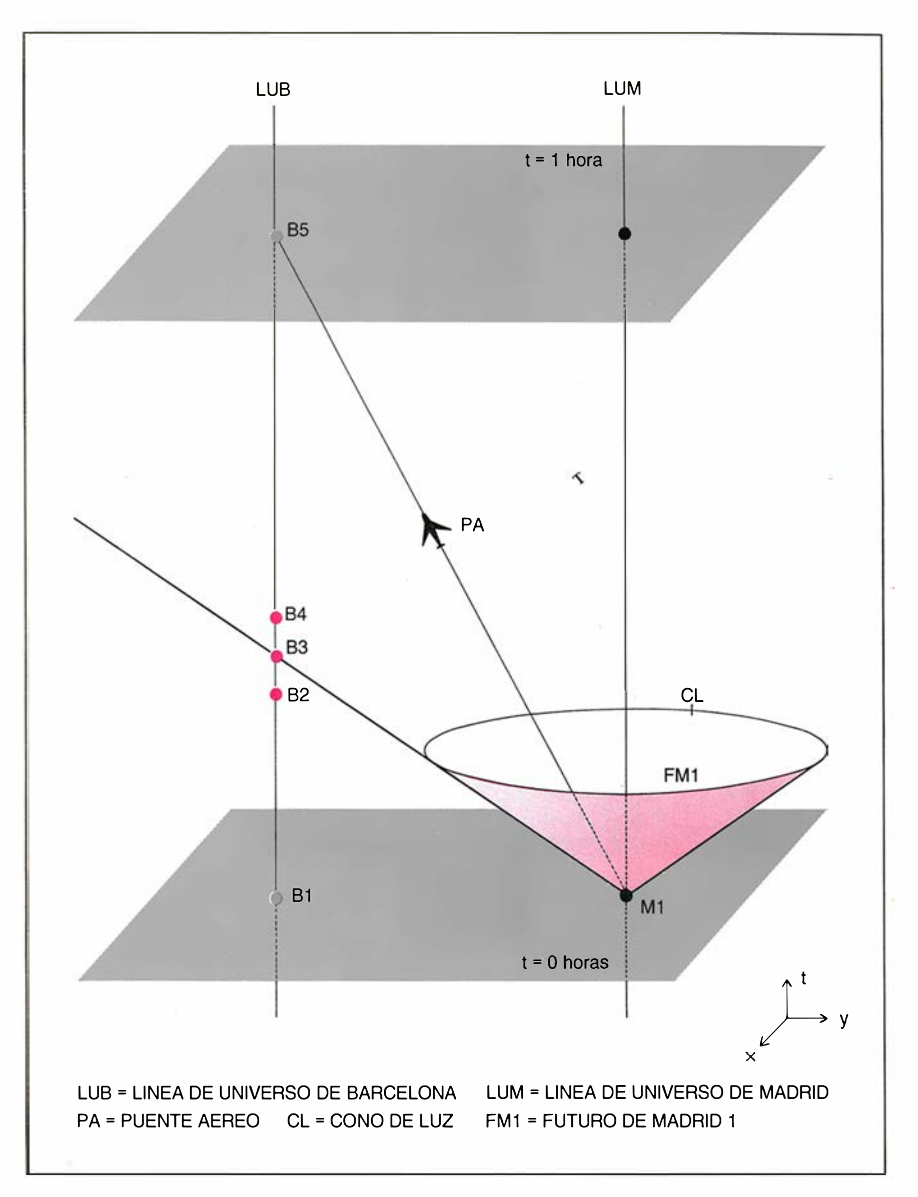
donde c denota la velocidad de la luz. Concluimos, pues, que, si un cuerpo de masa M está concentrado en un radio menor o igual que $2GM/c^2$, entonces nada, ni siquiera la luz, puede escapar de él. A tal radio se le denota r_s y se le denomina radio de Schwarzschild del cuerpo.

singularidad en RG quedaba establecido. Para estudiar el otro tipo, el de la gran explosión, debemos volver a pasearnos por la historia de la RG y de la cosmología.

La singularidad inicial

Einstein se percató muy pronto de que, dado que la interacción gravitatoria dominaba a grandes escalas, podía aplicar su nueva teoría de la RG para calcular el campo gravitatorio del universo y, obtener, por añadidura, la geometría del espacio-tiempo subyacente. Comenzaba así la cosmología relativista. Creíase en esas fechas que el universo era estático –o sea, más o menos inmutable–. Por eso, Einstein se llevó una gran sorpresa cuando observó que las ecuaciones de su teoría no admitían soluciones estáticas razonables. Para superar este problema, introdujo en sus ecuaciones un nuevo término, la constante cosmológica. Con esta constante, Einstein presentó un modelo cosmológico que fue seguido inmediatamente por otro modelo publicado por Willen de Sitter en 1917. Esta última solución contenía una singularidad ficticia (en realidad un horizonte), lo cual provocó cierta polémica entre Einstein y de Sitter, más o menos resuelta en el libro de Arthur Stanley Eddington *The Mathematical Theory of Relativity* (1923). En 1922 y 1924, el matemático soviético A. Friedman estudió los primeros modelos cosmológicos con singularidad real; su trabajo fue ampliado posteriormente por A.P. Robertson y más tarde por Walker. Estos modelos constituyen la base de la actual cosmología estándar.

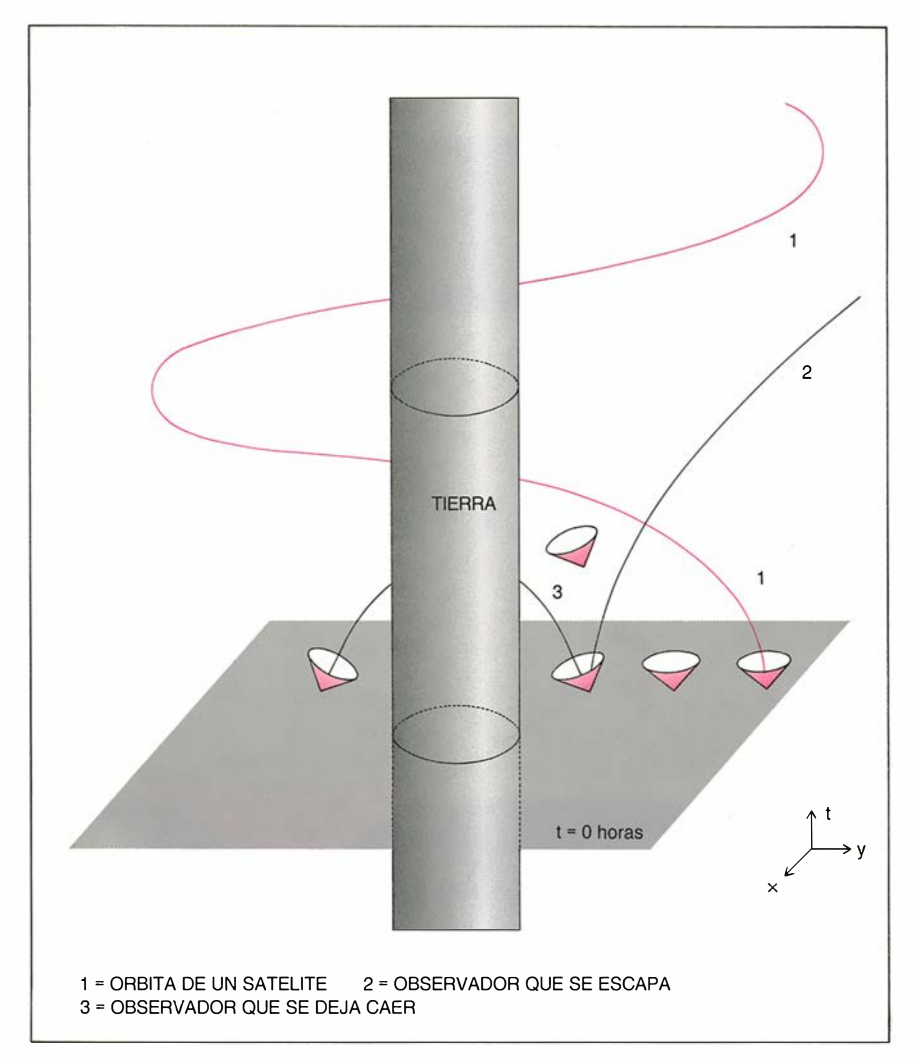
Desde un punto de vista experimental, el hecho más importante para la cosmología de los años treinta fue el descubrimiento, realizado por Edwin Hubble en 1929, del desplazamiento hacia el rojo de las rayas espectrales de las galaxias distantes y, como consecuencia, la relación lineal entre la distancia de tales galaxias a nosotros y su velocidad de alejamiento. El resultado de ello sería la aceptación de un universo en expansión. Hay corrimiento hacia el rojo en el modelo de de Sitter, aunque entonces la interpretación que se le daba era de otro tipo. Lemaître fue quien primero calculó los corrimientos al rojo en modelos similares a los de Friedman –no estáticos–, y su trabajo impresionó tanto a Eddington que éste le dio bastante publicidad. Se empezó a aceptar lo que hoy llamaríamos la interpretación normal de tales corrimientos al rojo, a pesar de que con ello algunos modelos cosmológicos presentaban una singularidad espacio-temporal.



2. DIAGRAMA ESPACIO-TEMPORAL Y ESTRUCTURA CAUSAL. En el espacio-tiempo plano, un observador en reposo describe una línea recta vertical. (El espacio es, en la ilustración, los planos bidimensionales horizontales. Las líneas rectas que describen Madrid (M) o Barcelona (B) se denominan líneas de universo, así como también la línea del observador que va en el puente aéreo de M1 a B5; el ángulo que forma ésta con los planos horizontales es una medida de la velocidad con la que se mueve con respecto a M o B, que están en reposo mutuo. Dado que la velocidad de la luz es la máxima velocidad de propagación, las líneas rectas descritas por los rayos luminosos emitidos desde M1 forman un cono de luz. Cualquier observador que haya pasado por M1 sólo puede viajar a los sucesos que están en el interior del cono de luz de M1. Todos los eventos en dicho interior están relacionados causalmente con M1 y al conjunto de todos ellos se le denomina futuro causal de M1.

Einstein comprendió que la constante cosmológica era más bien arbitraria, apreciación que compartió Richard Tolman, quien también presentó un modelo cosmológico no estático. Grosso modo, podemos resumir las posturas de los años treinta en dos clases: los que usaban la constante cosmológica y los que prescindían de ella –principalmente Einstein y Tolman– y estudiaban modelos en expansión, todos los cuales poseían una singularidad inicial, conocida hoy como "gran explosión". Los modelos estándar de Friedman-Robertson-Walker eran los únicos que satisfacían el principio cosmológico, que afirma que el universo es homogéneo (todos sus puntos son equivalentes) e isótropo (todas las direcciones espaciales son equivalentes). Este principio se basa en la observación de que el universo se nos aparece isótropo a nosotros y en la aceptación de que no ocupamos un lugar privilegiado, siguiendo las ideas iniciadas por Copérnico y Galileo.

Estos últimos modelos contienen una singularidad inicial que es, de hecho, una creación u origen de todo el universo. Antes de la gran explosión no había espacio-tiempo. Tan singu-



3. DIAGRAMA ESPACIO-TEMPORAL DEL CAMPO GRAVITATORIO TERRESTRE. Como en casos anteriores se suprime una dimensión espacial. La Tierra está, por consiguiente, representada por un disco en cada instante de tiempo y, en todo el espacio-tiempo, por un cilindro. Debido al campo gravitatorio terrestre, aquí muy exagerado, los rayos de luz se desvían, lo que produce una inclinación de los conos de luz, la cual es mayor cuanto menor sea la distancia al cilindro de universo de la Tierra. La línea de universo de un observador que se abandona a sí mismo ya no es una línea vertical, sino que se corta con la superficie de la Tierra. (En la figura trataríase de un observador que se deja caer.) También se representa la órbita –línea de universo– de un posible satélite que no se parece a una línea recta vertical, sino a una espiral. Esta órbita, proyectada en un plano t = constante nos da una elipse, la elipse de las órbitas keplerianas. Se ha representado también la línea de universo de un observador que se escapa: podría ser un cohete enviado a explorar el sistema solar.

lar propiedad no molestaba entonces a los físicos, ya que la inmensa mayoría pensaba que se debía a la excesiva simetría de las soluciones y que, en un pasado remoto –cerca de la gran explosión–, el estricto principio cosmológico no sería cierto, debido a la inevitable aparición de irregularidades.

A pesar de todo, el modelo de mayor éxito hasta los años sesenta era el del estado estable, formulado por Hermann Bondi, Fred Hoyle y otros. Este modelo estaba en expansión y no tenía singularidad, aunque necesitaba una continua creación de materia. Entretanto, fue George Gamow el primero en realizar un cálculo moderno de aparición de elementos en los modelos de Friedman; aunque conocía la existencia de la singularidad inicial, como otros muchos no le concedió importancia, pues pensaba que se debía a una fase previa de contracción del universo en la que toda la materia se había deshecho en sus constituyentes elementales (partículas elementales). Fue él, con sus colaboradores, quien predijo la existencia, si los modelos de gran explosión eran correctos, de una radiación de fondo, fósil de épocas antiguas superdensas y muy calientes. La detección de esta radiación de fondo por Arno A. Penzias y Robert W. Wilson, de los laboratorios Bell, en 1964 es, quizás, el descubrimiento de mayor relevancia para la cosmología y, ni que decir tiene, una indicación sólida de que el universo había pasado, en tiempos remotos, por una fase extremadamente compacta, muy densa y a alta temperatura. En posteriores observaciones se ha detectado que esta radiación de fondo es altamente isótropa, lo cual nos lleva a considerar seriamente la posibilidad de que el universo sea de tipo Friedman-Robertson-Walker ahora y lo haya sido en etapas anteriores. En consecuencia, quizá debamos aceptar la singularidad inicial.

Pero, al mismo tiempo que se detectó la radiación de fondo, Penrose estaba gestando, en 1964, el que sería un resultado de fundamental importancia y que abriría la puerta para el desarrollo de los teoremas.

Teoremas de singularidades

Ya sabemos que las singularidades aparecen en diferentes situaciones físicas dentro del marco de la RG. ¿Se deben estas singularidades a la idealización de los problemas y a la excesiva simetría impuesta para abordarlos? En el caso del colapso gravitatorio, supusimos que el cuerpo que se contraía poseía simetría esférica y que el colapso preservaba esta situación altamente perfecta. Tal posibilidad no es realista; bastan pequeñas desviaciones de la simetría esférica para provocar que no todo colapsara en el centro $r = 0$ y evitar así la singularidad. Además, aunque el universo parece isótropo, no lo es en puridad, habida cuenta de la existencia de irregularidades locales (las galaxias, por ejemplo).

Se necesitaba, pues, una respuesta a estas cuestiones para poder aceptar o desechar las singularidades. El primer problema, sin embargo, era cómo definir el concepto de singularidad. Porque las singularidades no forman parte del espacio-tiempo, los físicos no saben decir casi nada acerca de ellas. Podría pensarse que una singularidad viene definida por una zona donde alguna magnitud espacio-temporal se hace infinita, pero esto no es en absoluto cierto. Para Penrose, se dice que un espacio-tiempo contiene una singularidad cuando existen observadores que tienen un final repentino o un comienzo espontáneo. La razón para dar esta definición un tanto extraña y rebuscada es la lógica suposición de que las líneas de universo de cualquier observador no deberían tener un final ni un principio, por obvia comparación con el espacio-tiempo plano. Por otro lado, es trivial ver que cuando existen singularidades hay observadores con un final súbito o con un principio ex nihilo.

Penrose se planteó "la cuestión [...] de si la singularidad [en el colapso es-

telar] es, de hecho, simplemente una propiedad de la alta simetría... La materia colapsa radialmente al único punto central, así que la catástrofe resultante en el espacio-tiempo quizá no es sorprendente. La presencia de perturbaciones que destrozan la simetría esférica ¿no podría alterar la situación drásticamente?". Con este prefacio y el bagaje de refinadas técnicas matemáticas desarrolladas por él mismo, Penrose demostró la presencia de singularidades en el colapso estelar si se suponen ciertas condiciones razonables.

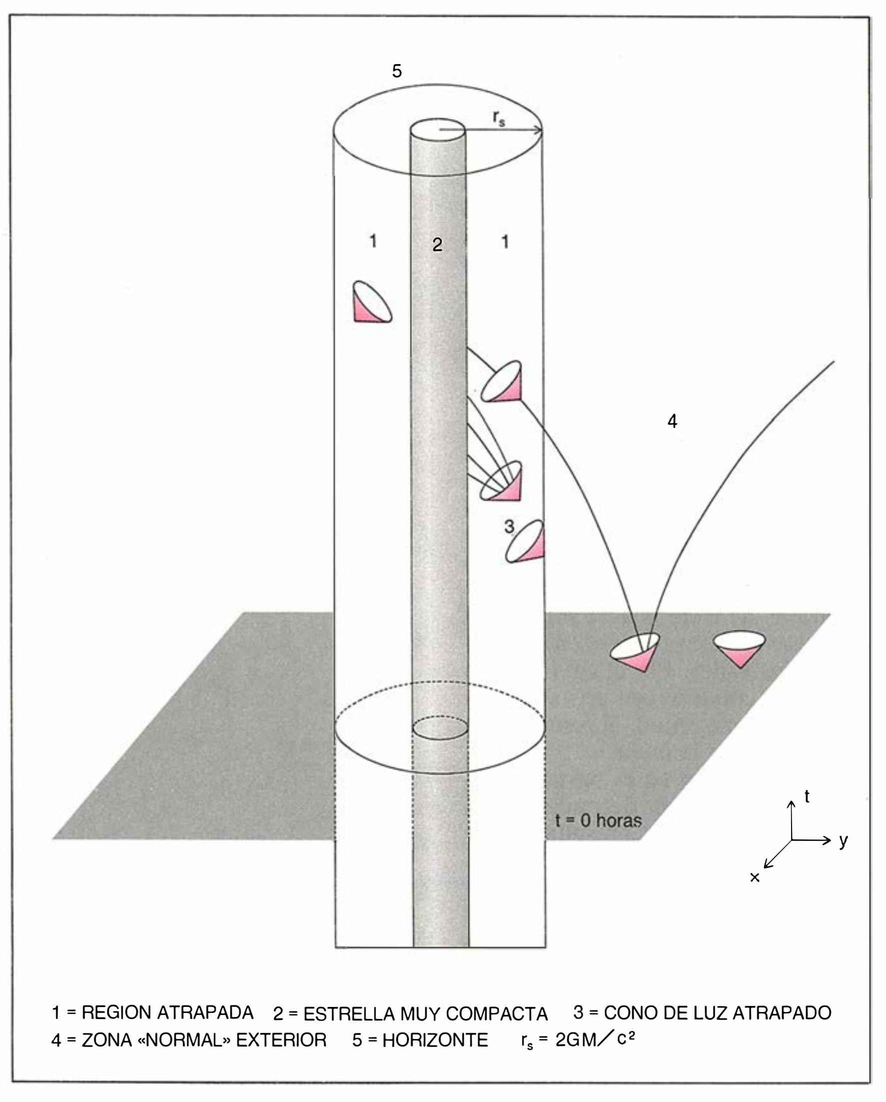
¿Cuáles son estas suposiciones? Esencialmente tres. Primero impuso una condición de energía, esto es, una condición según la cual la energía es siempre no negativa, lo que implica que la interacción gravitatoria sea siempre atractiva; suposición que está avalada por una firme base experimental. Segundo, adoptó una forma particular de condición de causalidad; en este caso concreto fue la condición según la cual todos los observadores deberían pasar, una y una sola vez, por el espacio definido en un instante de tiempo dado. Se puede obtener una idea intuitiva de esta condición observando que, en los diagramas de las figuras 2-5, cualquier línea de universo corta los planos t = constante una y una sola vez. Penrose impuso, por último, una condición inicial, a saber, la existencia de una superficie cerrada atrapada. Estas superficies son aquellas en las que su futuro causal está incluido en su interior. Con estas tres condiciones Penrose demostró de forma sencilla pero tajante que el espacio-tiempo debería contener líneas de universo incompletas, o sea, singularidades.

La importancia de este teorema es doble. Por una parte, indica que las desviaciones de la simetría esférica no impiden la formación de singularidades en el colapso gravitatorio, ya que según los resultados de Chandrasekhar y Oppenheimer parece claro que en tales situaciones existirán superficies atrapadas. Segundo, y quizá más importante, impulsó y estimuló a otros físicos a estudiar y extender estos resultados. Así fue como Geroch y, principalmente, Hawking publicaron una serie de resultados calificados como teoremas de singularidades. Hawking obtuvo teoremas aplicables a la cosmología y, en particular, a la singularidad inicial; todos presentaban el mismo esqueleto: una condición de energía, otra de causalidad y, finalmente, una condición inicial.

En 1970, Hawking y Penrose presentaron el que sería el teorema de singularidades por excelencia. En este teorema combinaron todos sus resultados anteriores y debilitaron las condiciones a exigir, de manera que incluía la mayor parte de las situaciones físicas de interés. Aunque posteriormente hubo mejoras de los teoremas por medio de los trabajos de Frank J. Tipler, Clarke y otros, el resultado esencial es el contenido en el teorema de Hawking y Penrose.

Podemos resumir este resultado de la siguiente forma. En el colapso gravitatorio, la singularidad es inevitable siempre que se formen regiones atrapadas. La aparición de éstas no está rigurosamente probada, pero todo parece indicar que así será a través de la extensión realizada de los trabajos de Chandarsekhar y Oppenheimer y, también, por la observación de la existencia de púlsares: estrellas de neutrones de masa enorme y radio muy pequeño. En el caso de la cosmología, por otra parte, hay dos posibilidades. Si el universo es cerrado –finito espacialmente– de nuevo parece inevitable la aparición de regiones atrapadas, como prueba la existencia, observada experimentalmente, de la radiación de fondo; la alta isotropía de ésta, a su vez, denuncia la posible existencia de una singularidad, que normalmente se interpreta como la gran explosión. Por el contrario, si el universo es abierto –espacialmente infinito–, la existencia de regiones atrapadas no parece obligada y hay casos con singularidad y casos sin ella.

Los teoremas de singularidades apoyan pues la idea de la consustan-



4. VISION INGENUA de la región atrapada de un objeto muy compacto. La situación es como la de la figura precedente, pero el objeto que crea el campo gravitatorio es muy compacto, de forma que su radio de Schwarzschild r_s es mayor que el radio real de la estrella. La corona cilíndrica es una región atrapada causalmente, ya que el futuro de cualquiera de sus sucesos está contenido en dicha región. De hecho, cualquier observador que se encuentre en la región atrapada caerá en la estrella. Todo ello se debe a la inclinación que sufren los conos de luz, como se muestra en la figura, que, en el interior de la región definida por el radio de Schwarzschild, apuntan hacia la estrella. En rigor, la situación física real no es ésta, sino la que se representa en la figura siguiente, ya que la propia materia de la estrella, una vez dentro de la región atrapada, cae hacia el centro de simetría.

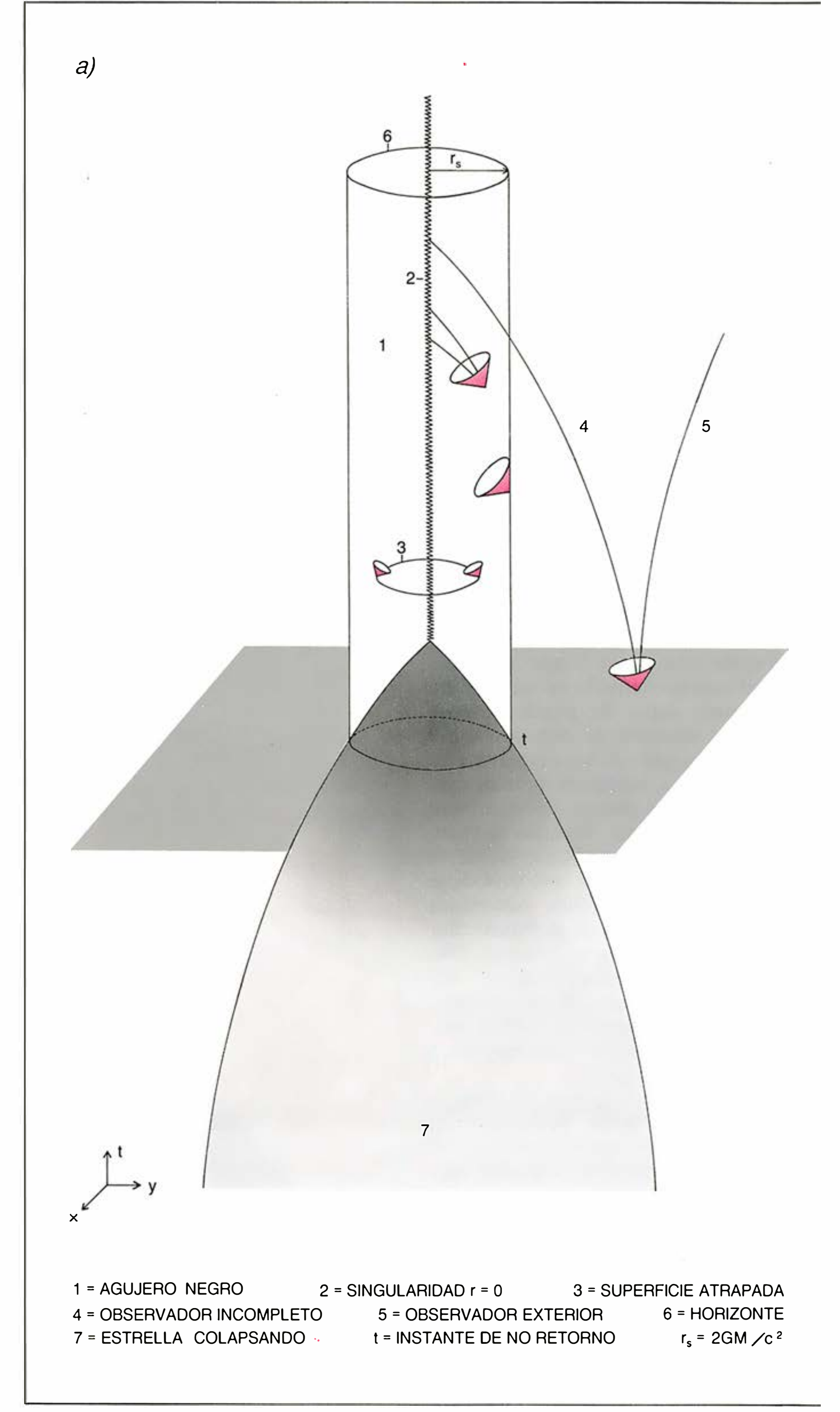
cialidad de éstas con la RG, siempre que supongamos condiciones físicamente realistas. Pero, ¿es esto realmente así? Aún no hemos hablado de un tercer tipo de singularidad que surge en la RG: la de la colisión de ondas gravitatorias.

Colisión de ondas gravitatorias

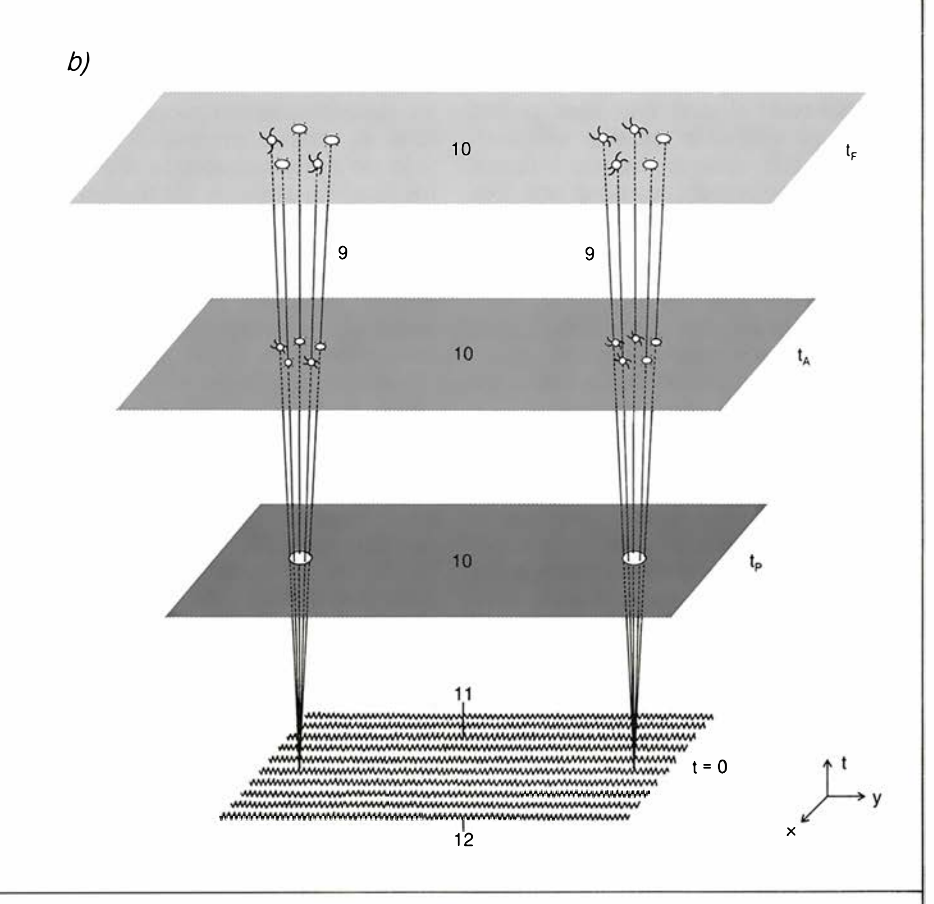
Otra de las predicciones de la RG –aún sin confirmación experimental definitiva– es la existencia de ondas de gravedad que se propagan, como las ondas electromagnéticas, a la velocidad de la luz. En la actualidad hay una gran inversión de medios tanto técnicos como humanos para la detección de tan débiles ondas. En todos los campos de la física, las ondas más simples son las ondas planas, que aparecen de forma aproximada en regiones suficientemente pequeñas muy alejadas de las fuentes que las originan. Uno de los primeros problemas que cualquier físico tratará de entender en cualquiera de estos campos es el de la colisión de ondas.

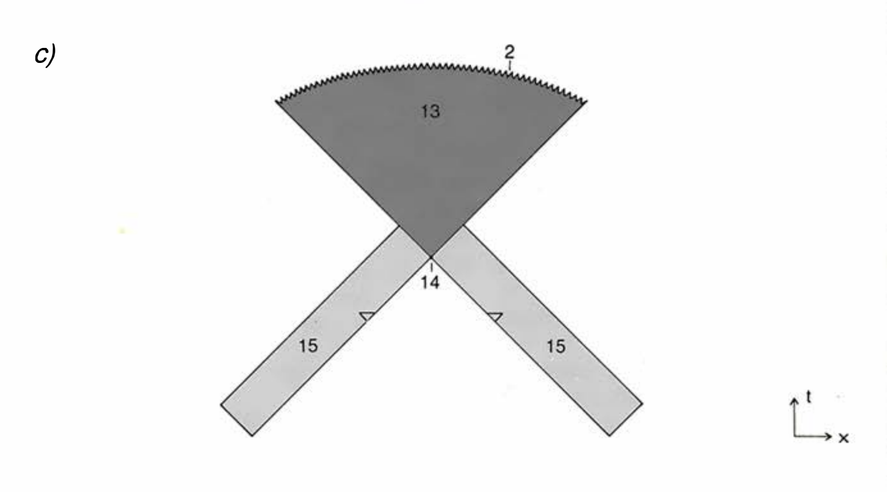
Podemos obtener una idea sencilla de una colisión de ondas pensando en lo que ocurre cuando lanzamos dos piedras a un lago, y en cómo las olas –ondas de agua– que estas piedras generan chocan e interfieren constructiva o destructivamente. En general, los casos de colisión estudiados en física son lineales –o sea, obedecen el llamado principio de superposición– y la onda resultante no es más que la simple suma de las ondas iniciales. Así, por ejemplo, si la cresta de una onda coincide en un punto con la cresta de la otra onda, el resultado será una cresta de altura la suma de las alturas de cada una de las crestas iniciales y lo mismo ocurre en un punto con dos valles. Por el contrario, si una cresta coincide con un valle, ambos de igual magnitud, el resultado es la anulación de la onda.

La situación en RG es bastante distinta. La interacción de ondas gravitatorias es no lineal. El resultado producido en su colisión difiere, por tanto, del antedicho y resulta inesperado. Las primeras soluciones obtenidas que describían colisión de ondas planas gravitatorias fueron presentadas por Khan y Penrose e, independientemente, por Szekeres en los primeros años de la década de los setenta. El resultado era la aparición, un tiempo finito después del primer instante de colisión, de una singularidad espacio-temporal que provocaba la catastrófica desaparición del espacio-tiempo. Además, esta singularidad era independiente de las intensidades de las ondas iniciales. Tan extraña situación podría no sorprendernos des-



5. LAS TRES SINGULARIDADES predichas por la relatividad general. En primer lugar, el colapso gravitatorio (*a*). Como en las figuras anteriores, se tiene una estrella, pero sufriendo un colapso debido a su propia atracción gravitatoria. Esta es tan fuerte, que el radio de la estrella decrece hasta que sobrepasa su radio de Schwarzschild. A partir de ahí, la estrella sigue contrayéndose inevitablemente hasta que toda su masa está concentrada en el centro $r = 0$, donde la densidad se hace infinita. En esta singularidad, todas las leyes de la física dejan de tener validez. La región atrapada que aquí se forma es un agujero negro, y ningún tipo de información puede escapar de dicha zona. Por tanto, un observador exterior nunca podrá ver la singularidad, salvo que sea tan osado como para entrar en el agujero negro, lo cual le forzaría a caer en la singularidad, sin poder comunicar sus descubrimientos al exterior. A pesar de todo, cualquier observador exterior seguirá sintiendo la atracción gravitatoria de la estrella colapsada. La segunda es la gran explosión o singularidad inicial (*b*). Desde el descubrimiento realizado por Hubble de que las galaxias se alejan de nosotros con





9 = LINEAS DE UNIVERSO DE LA MATERIA 10 = SUPERFICIE NO PLANA
11 = «SUPERFICIE» CON CURVATURA Y DENSIDAD INFINITAS 12 = SINGULARIDAD INICIAL
t_F = TIEMPO FUTURO t_A = TIEMPO ACTUAL t_P = TIEMPO PASADO
13 = REGION DE INTERACCION 14 = INSTANTE DE COLISION 15 = FRENTE DE ONDA

velocidad proporcional a su distancia, sabemos que el universo está en expansión. Suponiendo que no ocupamos ningún lugar privilegiado en el universo, cualquier otro observador típico verá lo mismo que nosotros. Entonces comprobaremos que la materia estaba, en el pasado, cada vez más comprimida. Un esquema se presenta en la figura, donde vemos que en el pasado debió haber un estado de densidad y curvatura infinita, a partir del cual se originó el universo. Aunque se han representado las secciones espaciales como planos, en realidad están curvados; la curvatura se hace cada vez mayor para tiempos remotos, hasta convertirse en infinita en la gran explosión. Por último, la tercera singularidad: colisión de ondas gravitatorias (*c*). En la figura se representan dos ondas planas viajando, a la velocidad de la luz, frente contra frente en la dirección espacial *x* –la única mostrada por mor de simplicidad–. Después de producirse la colisión hay una zona de interacción, y cuando ha transcurrido un tiempo finito, la mutua focalización –como la de las lupas– es tan grande que el espacio-tiempo desaparece y surge una nueva singularidad. En este espacio-tiempo no hay materia.

pués de toda la discusión anterior y de los teoremas de singularidades. Pero esta singularidad es profundamente distinta a las anteriores y, como vamos a ver, es mucho más problemática.

Las singularidades tradicionales de la física se consideran benignas por dos razones: representan las fuentes y no destruyen el espacio-tiempo. La gran explosión y los agujeros negros no cumplen la segunda propiedad, pero todavía mantienen la primera, ya que son singularidades donde se concentra la materia que es fuente del campo gravitatorio. Podríamos, pues, calificar de semibenignas a estas singularidades y, con una postura optimista, esperar que una adecuada teoría para la materia dentro de la RG suavice la semicatástrofe que producen. La singularidad del choque de ondas, contrariamente, ocurre en un espacio-tiempo vacío –sin materia–, y no hay posible interpretación optimista. Al no satisfacer ninguna de las propiedades de benignidad, la singularidad de la colisión de ondas podría ser calificada de maligna. Este problema ha sido expuesto claramente en diferentes ocasiones, en particular, por W.B. Bonnor.

El examen más pertinente del problema se debe a Tipler, quien considera –en contra de la opinión de Penrose y Khan– que la aparición de singularidades en la colisión de ondas planas depende crucialmente de la suposición de planitud de las ondas. (Una suposición que es altamente idealizada e irrealizable de forma exacta en la realidad.) Si este resultado fuera cierto, quizá podríamos recuperar el optimismo, pero la falta de soluciones exactas que describan la colisión de ondas no planas deja abierta la cuestión. Por lo demás, esta singularidad maligna reabre la cuestión de hasta qué punto la excesiva simetría es la causa de las singularidades y de si, en general, debemos culpar a la propia teoría de la RG, y a sus posibles fallos, de la aparición de singularidades; o, por el contrario, de si las singularidades son algo realmente existente. Pero, ¿cuál es la situación actual del problema de las singularidades?

Existencia de singularidades y RG

Pudiera creerse que las condiciones de energía –o sea, que la energía sea siempre positiva y, por tanto, la gravedad siempre atractiva– implican necesariamente las singularidades por medio de la RG. Esa idea, bastante extendida, es falsa. Según hemos visto, los teoremas de singularidades presuponen otras hipótesis, aparte de

una condición de energía; por ejemplo, las condiciones de causalidad. Por ello, unida a aquella idea falsa suele ir la creencia de que las condiciones de energía implican, a través de la RG, singularidades o violación de la causalidad. De nuevo, esto también es falso ya que, además de condiciones de energía y causalidad, los teoremas de singularidades dependen esencialmente de un tercer tipo de condiciones, a saber, las condiciones iniciales: en todos los casos, estas condiciones implican la existencia de regiones causalmente atrapadas, ya sea en el futuro o en el pasado. La conclusión que podemos obtener de la falsedad de estas dos ideas es directa, y un análisis objetivo y detallado del tema nos permite afirmar que la RG no implica, necesariamente, la existencia de singularidades.

Me gustaría abundar en este tema, pues la afirmación anterior podría calificarse de heterodoxa. Nada más lejos de la realidad. Existen muchísimas soluciones de las ecuaciones de Einstein de la RG libres de singularidades. Lo que sucede es que se comete el error generalizado de confundir dos cuestiones diferentes. La primera es la antedicha afirmación. La segunda es: ¿tiene singularidades nuestro universo? La respuesta a esta pregunta puede ser afirmativa, debido principalmente a la posible existencia de regiones atrapadas. Pero esto último es un hecho observacional y no forma parte del cuerpo teórico básico de la RG. La comprensión de esta diferencia debería eliminar el carácter heterodoxo de la negación de la implicación necesaria de las singularidades por la RG.

En el caso de la cosmología y la singularidad inicial, la aceptación de esas ideas falsas se basaba en el hecho de no haberse encontrado soluciones exactas cosmológicas –modelos cosmológicos– sin singularidades. Esto ha cambiado recientemente, debido a un modelo presentado por el autor que satisface las condiciones de energía y causalidad en sus versiones más fuertes, que verifica además otras muchas condiciones realistas, que tiene un carácter obviamente cosmológico y que no presenta singularidades. En este modelo, aunque la materia colapsa por su propia atracción, un gradiente de presión impide que colapse indefinidamente y la obliga a rebotar y comenzar a expandirse. Todo ello es una pura y simple consecuencia de la no existencia, en este modelo, de regiones atrapadas. La aparición de esta solución ha causado cierto revuelo y me gustaría señalar que simplemente nos indica, o mejor, recuerda, la veracidad de las anteriores consideraciones sobre la no implicación directa de las singularidades por la RG. Pero este modelo no puede dar una descripción completa del universo real, el cual muy bien podría haberse originado en una singularidad inicial, aunque, repito, a fuerza de parecer pesado, eso no es una consecuencia de la RG, sino de un dato observacional como la existencia e isotropía de la radiación de fondo y de posibles regiones atrapadas.

Todo lo anterior nos conduce, a mi parecer, a una nueva pregunta. Aceptando las dos posibilidades reales, en pie de igualdad aproximadamente, de soluciones con y sin singularidades, y aceptando asimismo que nuestro universo parece haberse originado en una singularidad inicial, ¿por qué el universo ha escogido un origen tan singular? Esta cuestión podría parecer más filosófica que científica, pero la cosmología se plantea preguntas similares muy a menudo. Baste como botón de muestra lo siguiente: el hecho de que nuestro universo sea bastante homogéneo e isótropo en la actualidad sólo parece poder explicarse con la suposición de condiciones iniciales muy especiales –y que forman una parte despreciable dentro del conjunto de todas las posibles condiciones iniciales–, lo cual ha llevado al estudio de modelos inflacionarios, por ejemplo. De la misma forma, podríamos preguntarnos el porqué de la elección del universo de condiciones iniciales especiales –o, al menos, en pie de igualdad con otras– tales como la gran explosión. No sé si la respuesta a esta cuestión está a nuestro alcance hoy en día, pero al menos deberíamos ser lo suficientemente honestos como para planteárnosla.

Cambiando un poco el discurso, es necesario plasmar aquí las posturas de los relativistas con respecto a las singularidades. Como norma general, todos aceptan el hecho de que las singularidades indican un límite de validez para la teoría de la RG. En otras palabras, a las altísimas energías a las que aparecen las singularidades, la RG no es una teoría fiable –no hemos podido comprobarla experimentalmente bajo tales condiciones extremas– y, por consiguiente, una teoría más adecuada y general sería necesaria. En este sentido, la postura absolutamente mayoritaria es aquella según la cual se espera que la combinación de los efectos cuánticos con los gravitatorios nos proporcione una solución satisfactoria del problema. Desafortunadamente, la teoría cuántica de la gravedad parece resistirse, y todos los intentos para su consecución han sido infructuosos hasta la fecha. Como siempre, la solución y la comprensión finales de estas cuestiones sólo el tiempo y el esfuerzo nos las darán.

Para acabar este breve repaso por el mundo de las singularidades en la RG, desearía plasmar mi personal opinión de que las singularidades plantean un problema abierto aún en estos días, y a pesar de los teoremas de singularidades. Así, quisiera terminar enunciando una pequeña serie de preguntas que, a mi entender, necesitan urgente respuesta. ¿Cómo resolver el problema de la singularidad maligna en la colisión de ondas gravitatorias? ¿Qué tipo de solución es más genérica en RG, la que tenga o la que no tenga una singularidad? Si ambos tipos son igualmente genéricos, aproximadamente, ¿bajo qué razonamiento aceptar o desechar unas u otras? En este sentido, ¿por qué el universo habría elegido tener su origen en una singularidad inicial? ¿Es posible probar un teorema de no singularidades, estableciendo bajo qué condiciones físicamente realistas una solución no tendrá singularidades?

Estas y otras muchas preguntas que no incluimos en aras de la brevedad no tienen una respuesta satisfactoria todavía. Y, lo que a mi parecer es peor, no se les está dedicando la atención que merecen. Cualquier aproximación a una respuesta a esas cuestiones nos ayudará indudablemente a la definitiva comprensión de las singularidades en RG y, en particular, a profundizar en el posible origen –y el posible destino– de la materia.

BIBLIOGRAFIA COMPLEMENTARIA

THE SINGULARITIES OF GRAVITATIONAL COLLAPSE AND COSMOLOGY. S.W. Hawking y R. Penrose en *Proc. R. Soc. London, Ser. A* **314**; pág. 529; 1970.

THE LARGE SCALE STRUCTURE OF SPACE-TIME. S.W. Hawking y G.F.R. Ellis. Cambridge University Press; Cambridge, 1973.

GLOBAL STRUCTURE OF SPACETIMES. R. Geroch y G.T. Horowitz en *General Relativity. An Einstein centenary survey*, pág. 212, ed. S.W. Hawking y W. Israel, Cambridge University Press, Cambridge, 1979.

SINGULARITIES FROM COLLIDING PLANE GRAVITATIONAL WAVES. F.J. Tipler. Phys. Rev. D **22**, 2929, 1980.

SINGULARITIES AND HORIZONS –A REVIEW ARTICLE. F.J. Tipler, C.J.S. Clarke y G.F.R. Ellis en *General Relativity and Gravitation. One Hundred Years After the Birth of Albert Einstein*, A. Held. Plenum Press; Nueva York, 1980.

GLOBALLY REGULAR SOLUTIONS OF EINSTEIN'S EQUATIONS. W.B. Bonnor. *Gen. Rel. Grav.* **14**, 807, 1982.

NEW CLASS OF INHOMOGENEOUS COSMOLOGICAL PERFECT-FLUID SOLUTIONS WITHOUT BIG-BANG SINGULARITY. J.M.M. Senovilla en *Physical Review Letters*, vol. 64, pág. 2219; 1990.

Ciencia y sociedad

Brazo de tierra

En 1751 un campesino francés soñaba construir un túnel submarino que le llevara hasta Inglaterra. El proyecto recibió la aprobación, en 1802, del mismo Napoleón. Volvióse a proponer en 1830, entonces por un ingeniero galo. Luego, en 1975. Por fin, el 30 de octubre de 1990 se completó el primero de los tres túneles que, bajo el mar, unirán Gran Bretaña con Francia. Desde hace unos 8000 años, durante la postrera glaciación, el agua cubre la franja de tierra que unía ambos países.

Esta proeza de la ingeniería se debe a Eurotunnel, compañía franco-británica constituida hace cinco años. Pero las cosas no se han sucedido de acuerdo con el estudio del proyecto de operación. La empresa invirtió, de partida, 9000 millones de dólares. A comienzos de 1990 el proyecto estuvo a punto de suspenderse cuando se solicitó a los accionistas un nuevo desembolso de 5000 millones de dólares. Tampoco se esperaban tantos obstáculos técnicos como han ido apareciendo a lo largo de la perforación de los 52,5 kilómetros de arcilla que separan las terminales a cada lado del Pas de Calais, la zona más estrecha del canal de la Mancha.

El empeño conjunto de Eurotunnel y de Transmanche Link, esta segunda asociada a aquélla y que integra un consorcio de compañías constructoras británicas y francesas, convierte el proyecto en un hito de la ingeniería civil del siglo xx. Más de 14.000 trabajadores remueven varios millones de toneladas métricas de tierra y roca. Manejan perforadoras gigantes, en condiciones de alta humedad y presión elevada. Reparan transportadores atascados y utilizan elementos nuevos sin previo ensayo de los mismos. Tapizan trozos de túnel entre chorros continuos de agua que mojan las líneas de conducción de alto voltaje. Hasta ahora ha habido que lamentar nueve bajas.

Se ha construido el túnel de servicio para automóviles. La compañía ha excavado el 70 por ciento de los otros dos túneles que acogerán a los trenes de gran velocidad. Si los trabajos avanzan al ritmo actual, la perforación acabará a finales del año en curso. En junio de 1993, pasajeros y vehículos podrán ir de Calais a Folkestone en 30 minutos.

Desde el comienzo, el proyecto ha sido un híbrido curioso de viejas ideas e innovaciones arriesgadas, de la conjunción de prácticas tradicionales con la ingeniería más avanzada. Así, aunque midieron la separación exacta entre orillas con radar, los topógrafos calcularon sus posiciones por triangulación y señalaron sus puntos de referencia con plomadas.

Bajo tierra, los técnicos no tuvieron otra opción que guiarse por su propio criterio cuando se vieron frente a los mandos de las perforadoras de los túneles. Cada máquina se mantiene en la dirección deseada ayudada por el haz de un láser. A medida que las perforadoras avanzan desde los dos extremos del túnel, las computadoras incorporadas calculan la posición del haz y el ángulo, comparando su avance con un mapa tridimensional del túnel almacenado en su memoria. De acuerdo con ello, imparten las instrucciones al operador de la máquina.

A cada lado del canal, y en cada túnel, se colocaron dos máquinas de perforación, una para excavar el interior del túnel y la otra las terminales del tren. Los trabajos del lado británico van por delante del francés, porque la perforación realizada durante los años setenta dejaron dos túneles de acceso en los acantilados de Shakespeare, al sur de Dover.

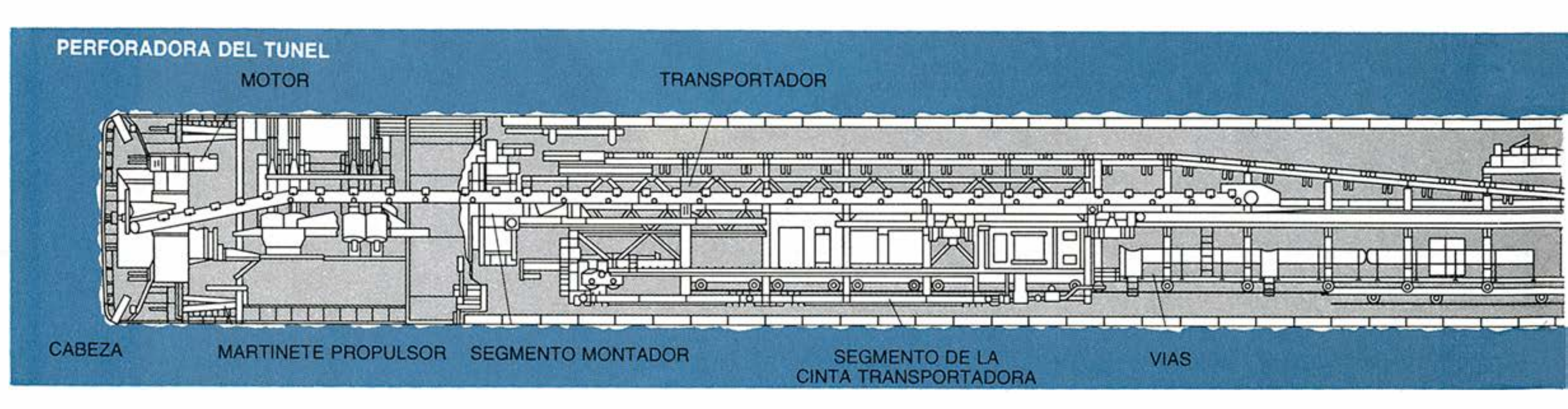
Las máquinas perforadoras son auténticos monstruos cilíndricos. Parecen gusanos gigantescos extraídos de un relato de fantasía científica que roen la pared con picos de tungsteno y discos cortantes montados en su cabeza giratoria y expulsan luego el material a través de bombas y cintas transportadoras.

El problema logístico más complejo fue el del transporte de los trabajadores, el transporte del equipo y de los suministros a los puntos de perforación, así como la recogida y acarreo de material al exterior. En todo su recorrido de 37,9 kilómetros bajo el agua el túnel carece de pozos de acceso. La única forma posible de entrar y salir es por las dos bocas principales.

El sistema de transporte ha acabado por convertirse en una red de más de 100 locomotoras que trabajan en 190 kilómetros de vía. Se alimentan de un tendido eléctrico y, a intermitencias, con baterías. (No se utilizan locomotoras diesel por los problemas de ventilación que encierran.) Las baterías impulsan las locomotoras un kilómetro, ayudándoles a cubrir los últimos centenares de metros de túnel donde el desarrollo de los trabajos torna peligrosa la instalación del tendido eléctrico.

En el lado inglés, hubo problema desde el comienzo con los trenes. La máquina de perforación pesa 1500 toneladas y ocupa unos 300 metros. Se montó pieza a pieza en los túneles de acceso bajo el acantilado de Shakespeare. Los trenes acarreaban las piezas nuevas a medida que la máquina iba abriéndose su propio espacio. Pero los túneles de acceso resultaron demasiado estrechos y las piezas de la máquina perforadora demasiado grandes: no cabía el tendido eléctrico.

Por eso, los ingenieros tuvieron que trabajar con trenes propulsados con baterías. Se sustituía fuera las baterías agotadas por otras cargadas que servían para que el tren recorriera 1,5 kilómetros túnel adentro. Por culpa de tan complicada logística, la perforadora solía parar con excesiva frecuencia, a la espera de alguna pieza o de algún recambio. Se tardó seis meses en ensamblar la máquina entera.

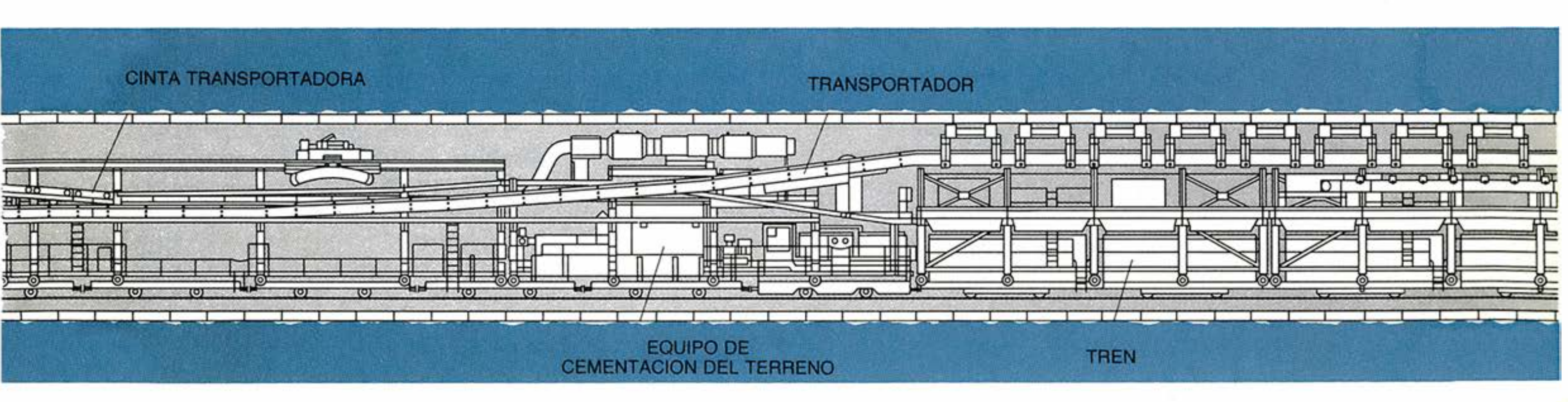


1. LA PERFORACION se realiza con una máquina de 8,4 metros que excava la creta del canal de la Mancha.

A pesar de estas dificultades, al comenzar la perforación regular, los ingenieros británicos se encontraron que la excavación avanzaba con facilidad, ya que la mayor parte del túnel se horadaba en una marga de creta, roca blanda e impermeable al agua que conserva su forma durante muchos años después de ser perforada. El estrato era grueso y casi horizontal, lo cual permitió al equipo de perforación seguir una trayectoria casi rectilínea.

Merced a esas condiciones, los ingenieros británicos pudieron utilizar un método de revestimiento de los túneles rapidísimo. En cuanto la cabeza avanzaba 1,5 metros, en el otro extremo de la máquina los trabajadores colocaban bloques de hormigón prefabricado formando un anillo. Los bloques se mantenían fijos en su sitio con una pequeña cuña que se colocaba en la parte superior del anillo. En cuanto se inyectaba entre las paredes del túnel y el anillo un lechado de hormigón los bloques creaban un arco doble clásico; la presión permanente de la arcilla se repartía lateralmente por las paredes del hormigón.

Los británicos tuvieron algunos problemas con capas húmedas entre el primer y el quinto kilómetro. En el túnel de servicio, se habían preparado para lo peor; lo revestían con una cubierta de hierro prefabricada, más hermética y, también, más difícil de instalar y más cara. En los túneles de trenes, el revestimiento se hacía con hormigón prefabricado, que goteaba como un tamiz. Los trenes eléctricos se movían sobre rieles sumergidos en



agua, causa de continuas interrupciones de los sistemas eléctricos. El émbolo percutor usado por la perforadora estaba en continuo peligro de romperse. Sobre los trabajadores llovía con intensidad. "Parecía que estábamos bajo una granizada", confiesa Andrew Copperwheat, ingeniero de Hunslet GMT Ltd., empresa suministradora de las locomotoras.

John R. J. King, director de túnel, comenta: "Desconocíamos las pérdidas que experimentan los sistemas eléctricos en condiciones de alta salinidad. Tardamos 12 meses en superar las dificultades técnicas". La compañía [Transmanche] ha optado por convertir diez de sus locomotoras eléctricas en diesel.

Tampoco las excavadoras francesas tuvieron mejor suerte. A ellos, además, les perjudicó la geología. La capa de creta, al llegar a la costa francesa se hunde; por tanto, antes de alcanzar el estrato impermeable, hubo que perforar en creta gris, material quebradizo muy permeable al agua. Para montar la perforadora, excavaron un gigantesco pozo de 55 metros de diámetro y 66 metros de fondo en Sangatte. En cuanto se impermeabilizó el recinto, los ingenieros instalaron cuatro grúas que bajaban los componentes de la inmensa oruga.

Los franceses esperaban encontrar mucha agua y encargaron una perforadora capaz de resistir hasta diez atmósferas de presión, muy por encima de las tres o cuatro atmósferas habituales en las perforaciones bajo agua. La máquina posee una cabeza hermética, sellada, de la que se bombea el material arrancado mediante un tornillo de Arquímedes, un eje rotativo con estrías espirales, y lo envía desde la cámara presurizada hacia la cinta transportadora en la zona de trabajo. La máquina también contiene una zona impermeabilizada donde los trabajadores montan el revestimiento del túnel. A diferencia del revestimiento usado en el lado británico, el francés se sella laboriosamente y se impermeabiliza con relleno de caucho.

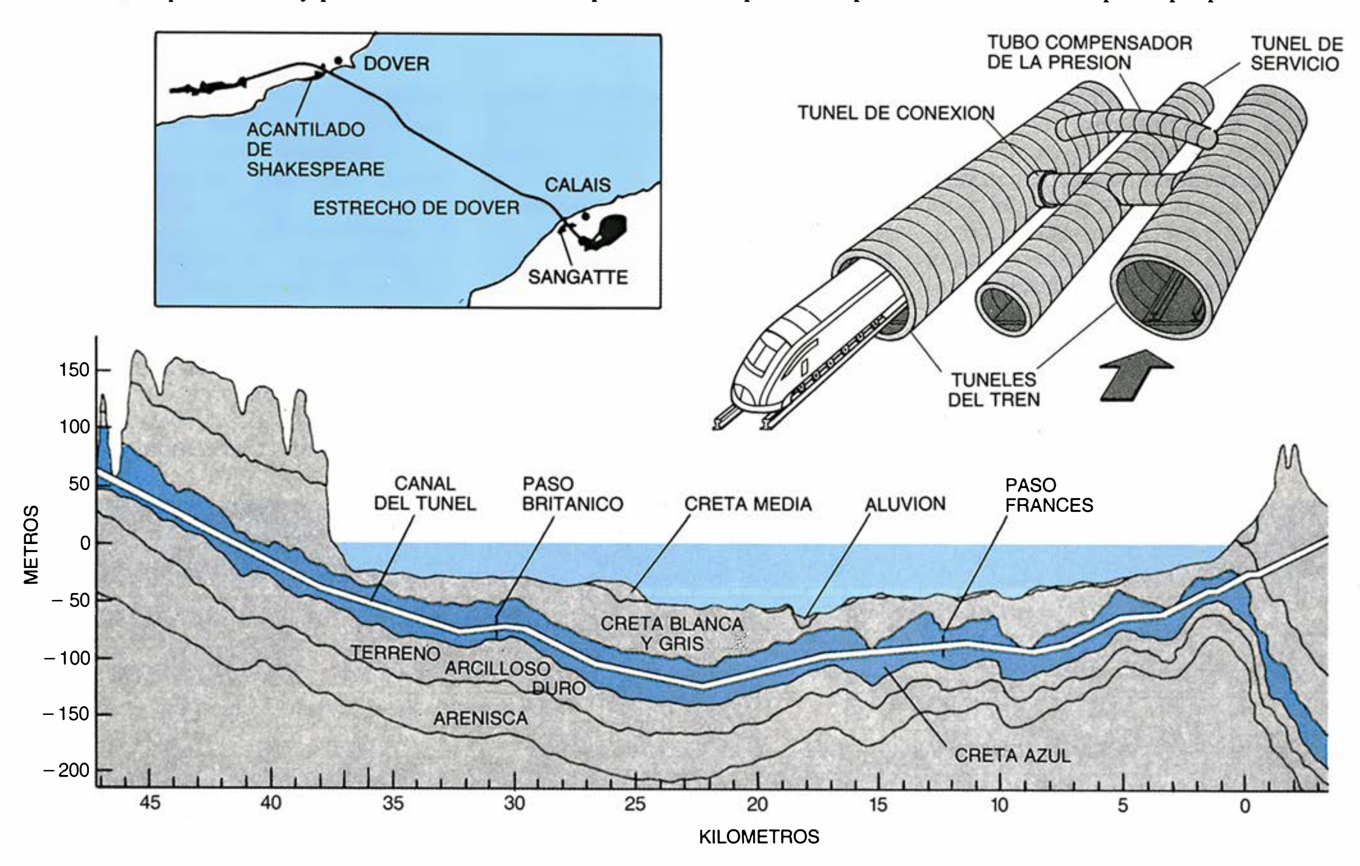
Durante el primer año de trabajo, los ingenieros franceses avanzaron sólo un kilómetro. Después de excavar durante los primeros 500 metros de creta gris, comenzó la marga de creta mucho más cómoda. Las perforadoras se vieron obligadas, algunas veces, a ondular su recorrido en el estrato cortando la capa de marga permeable y atravesando varias fisuras.

Si fue lento el comienzo a ambos lados del canal, el ritmo de la perforación se avivó después del primer año. Los británicos evitaron los problemas del agua en los túneles del tren mejorando su técnica de revestimiento. A mediados de 1990 se perforaban semanalmente más de 300 metros de túnel. También cambió la suerte francesa. Sus perforadoras pronto superaron las previsiones, avanzando 290 metros cada semana. A mediados de noviembre faltaban 14,6 kilómetros para perforar en el túnel norte y 18 kilómetros del túnel sur. A pesar de todos los problemas parece razonable prever que antes del verano termine la excavación.

Ambos equipos se hallaron ante un formidable reto técnico: la excavación de dos corredores, de unos 60 metros de alto y 20 de ancho, perpendiculares a los dos túneles de trenes. Se les destina para servir de desvío y cambio de vía de un túnel a otro. Permitirán cerrar un tramo de túnel para proceder a los trabajos de mantenimiento e inspección sin interrumpir el servicio. La excavación de esos pasos es delicada. Primero se perforan pequeños túneles, revestidos provisionalmente de hormigón. Se horadan luego los túneles creando, poco a poco, un boquete mayor.

Ahora que Transmanche Link y Eurotunnel conocen la geología del canal y que lo peor de la excavación está concluido, los ingenieros acometerán la segunda fase del proyecto, la instalación del material fijo: los raíles, los equipos de ventilación, las luces de alta intensidad y el conducto refrigerador.

La seguridad de los futuros pasajeros ha impuesto medidas muy estrictas de control. Entre las grandes amenazas hay que citar la del fuego. Una vía de escape la proporciona el



túnel de servicio, que está conectado con los túneles de trenes mediante múltiples conductos de ventilación y estaciones de bombeo.

Vector neural

Mediante la transformación de un virus letal para que trabaje en su beneficio, los neurólogos esperan acometer la terapia génica en el cerebro. El virus (herpes simplex) es atraído naturalmente hacia las células nerviosas, lo que le convierte en un portador ideal de genes extraños que pueden corregir enfermedades del sistema nervioso. Aunque se ha publicado poco sobre esas posibilidades del herpes simplex, los investigadores están apresurándose para aumentar los beneficios del virus y mantener a raya sus peligros.

"Es algo excitante. Los neurobiólogos se mueren por aplicar la terapia génica en el cerebro", dice Joseph C. Glorioso, catedrático de genética molecular y bioquímica en la Universidad de Pittsburgh. "Y pienso que ésta es la única manera de hacerlo." Aunque las pruebas en humanos pueden tardar todavía años, la técnica abriría el camino al tratamiento de las enfermedades de Alzheimer, Parkinson, Huntington y otras.

Algunos pioneros de la terapia génica están insertando genes en las células con retrovirus, que puede incorporar el gen corrector en el ADN de la célula que infecta. Para que el virus consiga esta integración, la célula infectada tiene que replicarse. Ahora bien, las neuronas maduras ya no se dividen; los retrovirus no pueden, por consiguiente, actuar en el sistema nervioso.

En cambio, el virus herpes simplex no sólo infecta neuronas, sino que expresa genes extraños incorporados en su ADN, sin requerir que la propia neurona se reproduzca. El virus permanece aquí en estado de reposo: no se replica ni mata a las células nerviosas, a menos que sea excitado por el estrés o algún factor ambiental. Aunque el 90 por ciento de los seres humanos son portadores de herpes simplex, este virus reside de forma habitual y tranquilamente en neuronas periféricas, no en las del cerebro, explica Xandra O. Breakefield, de la Facultad de Medicina de Harvard. "Es el único virus que sepamos que pueda introducirse en una neurona y permanecer latente."

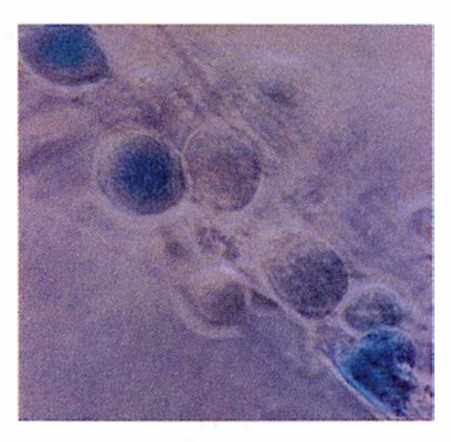
La latencia es crucial, porque aun cuando el virus no se reproduzca puede expresar algunos genes, los que son insertados o controlados por los investigadores. "El fundamento está ahí", dice Breakefield, cuyo labora-



2. SECCION recién excavada del túnel del tren.



3. ZONA de acceso en Sangatte, Francia.



4. NEURONAS infectadas por el herpesvirus; expresan un gen que hace que las células infectadas se vuelvan azules. (Cortesía de Alfred I. Geller.)

torio ha conseguido expresar con éxito genes que pueden teñir de azul la neurona, una prueba de color que confirma que el virus funciona según lo proyectado. "Ahora debemos insertar otro gen más importante."

El virus herpes simplex posee más ventajas. Merced a su talla (está compuesto de unos 70 genes) puede insertar en las neuronas genes extraños grandes, o más de un gen pequeño. Algunas cepas del virus pueden penetrar en el sistema nervioso periférico y viajar hasta el nervioso central, lo que obvia la necesidad de inyectar el virus directamente en el cerebro.

Pero sigue siendo potencialmente mortífero. "El tema de la terapia génica ha de abordarse con suma prudencia", advierte Breakefield. "Estamos hablando de un virus que puede matar el cerebro." En cualquier momento (que los científicos no pueden saber por anticipado) el virus se tornará activo, se multiplicará y destruirá la neurona.

Para contrarrestar su toxicidad, algunos laboratorios emplean cepas debilitadas o defectuosas del virus que no siempre son letales. (Sin embargo, algunas de éstas producen infecciones cerebrales en los animales de laboratorio.) Los científicos también alteran el virus de manera que ya no se replique.

Howard Federoff, de la Facultad de Medicina Albert Einstein, y sus colegas suprimieron más del 80 por ciento del ADN vírico, produciendo un virus inocuo, incapaz de reproducirse. Federoff insertó un promotor, que activa determinados genes, específico para las neuronas. Por tanto, los genes se expresan solamente en las células nerviosas.

Federoff cortó demasiado el virus para hacerlo benigno, argumentan algunos, y no expresará genes en animales vivos durante un período prolongado. El virus tiene que parecerse a su estado original cuanto sea posible si ha de servir para la terapia génica, dice Breakefield. Pero Alfred I. Geller, del Instituto del Cáncer Dana-Farber, que colaboró con Federoff, señala que su forma del virus expresa genes en ratas adultas durante por lo menos seis semanas.

Otros laboratorios emplean un sistema distinto para evitar la replicación del virus. No lo alteran mucho; prefieren reprimir su replicación y su capacidad de reactivación mediante la eliminación de dos genes necesarios para estas funciones, revela Jack G. Stevens, de la Universidad de California en Los Angeles. "Puedes hacer que este agente exprese cualquier gen que desees", afirma Stevens, quien está estudiando la introducción de factores de crecimiento nervioso en las neuronas a través de la lanzadera vírica.

En Pittsburgh, Glorioso emplea también un sistema represor. Explica que dentro de pocos meses utilizará el virus como transportador del gen de la tirosina hidroxilasa (que transcribe una enzima que conduce a la producción del neurotransmisor dopamina) para curar la enfermedad de Parkinson en primates. (El temblor y la rigidez muscular que experimentan las personas afectadas por la enfermedad de Parkinson se debe, al menos en parte, a la insuficiencia de dopamina.) Algunos investigadores advierten que la expresión génica a largo plazo que Glorioso ha obtenido podría interpretarse como una infección; Glorioso responde que esa acusación carece de fundamento.

Aparte de su potencialidad de replicación o reactivación, otros aspectos del virus pueden resultar peligrosos para las células. "Parte de su cubierta proteínica es tóxica", recuerda Breakefield, aunque añade que sólo las neuronas en placas de cultivo parecen ser afectadas por la cápside vírica. Los virus pueden sufrir mutaciones, y "muchas mutaciones son tóxicas", insiste Paul A. Johnson, de la Universidad de California en San Diego. Johnson está construyendo diferentes mutantes de herpes en un esfuerzo para reducir la citotoxicidad.

Púlsares, nuevos y viejos

Astrónomos de varios países han detectado en el centro de nuestra galaxia un voraz objeto celeste, un púlsar, que se está alimentando, tal parece, de su estrella acompañante. El gas absorbido sigue una trayectoria espiral que cela de vez en cuando las señales de radio. Esta observación proporciona una valiosa información que sirve para contrastar la teoría sobre la evolución de una nueva clase de púlsares.

Andrew G. Lyne, de la Universidad de Manchester, Richard N. Manchester, del Complejo Nacional de Telescopios de Australia, y sus colegas observaron el sistema binario apuntando sus radiotelescopios al cúmulo globular Terzan 5, situado a unos 21.000 años luz de la Tierra. Descubrieron allí un púlsar, clasificado según sus coordenadas PSR 1744-24A, que da una vuelta cada 11,56 milisegundos.

Los púlsares que giran con tamaña celeridad podrían seguir una evolución diferente de la tomada por otros púlsares. La mayoría de estos objetos giran con períodos superiores a la décima de segundo. Se cree que guardan una relación de origen directa con el proceso que forma las supernovas. La materia de la parte exterior de la estrella explota y crea una nebulosa. La materia de las zonas internas se colapsa hasta adquirir una densidad tal que se promueve la combinación de los electrones con los protones para formar neutrones. La rotación de la estrella de neutrones insta su característica emisión de ondas de radio al espacio.

Los púlsares no se ajustan a este modelo. Los astrónomos sospechan que comienzan a existir integrados en una estrella binaria. Se desarrollan lentamente, hurtando materia de su estrella compañera. A medida que su masa crece, aumenta su velocidad de rotación.

En 1982, Donald C. Backer, de la Universidad de California en Berkeley, Shrinivas R. Kulkarni, del Instituto de Tecnología de California, y sus colegas recabaron pruebas de ese cuadro cósmico. Descubrieron el púlsar que gira más rápido de cuantos se conocen, con un período de 1,56 milisegundos, unas cien veces más rápido que el púlsar medio. El período y el débil campo magnético abonaban la hipótesis de una gestación lenta, que se prolongaba decenas de millones de años. En 1988, Andrew S. Fruchter y sus colegas, de la Universidad de Princeton, encontraron un sistema de púlsar binario que se eclipsaba. En el sistema en cuestión, constituido por una estrella en órbita alrededor del púlsar, aquélla eclipsa periódicamente la señal de éste.

El púlsar encontrado por Lyne y sus colegas podría ser el eslabón perdido de la cadena evolutiva de estos objetos. Los sucesivos eclipses y el lento período de rotación del PSR 1744-24A parecen corresponder a un sistema más joven que el descubierto por Fruchter.

Según Lyne, el nuevo púlsar fue antes moribunda estrella de neutrones, con pausada rotación y escasa radiación. Pero no es improbable que pueda darse en los cúmulos globulares algún tipo de colisión. La gravedad de la estrella de neutrones capturó una estrella pequeña y empezó a engullir el material más superficial de su compañera, cuando ésta, en su evolución, llegó al estadio de gigante. Estos son sistemas binarios de rayos X de poca masa.

Conforme la estrella de neutrones va ganando masa, empieza a girar cada vez con más celeridad, con una rotación que le facilita adquirir una nueva vida de púlsar. Como todo púlsar, no le basta con su propia gravedad para mantenerse. La rápida rotación produce la emisión de radiación muy energética, capaz de vaporizar materia. Segun Lyne, "estamos observando el proceso de cocción del compañero del púlsar". El gas, al desprenderse, sigue un camino espiral que periódicamente interrumpe las señales de radio que llegan a la Tierra. El material de la compañera, que según parece es una enana blanca, engloba a veces al púlsar. Este proceso se sigue dando hasta que su compañera, la enana blanca, está próxima a su fin.

Se habla de los púlsares binarios como fuentes ideales para someter a prueba a la relatividad general. Los astrónomos creen que los sistemas binarios pierden energía en forma de ondas gravitacionales y que el ligero cambio en la órbita y en el período debe ser detectable. Por desgracia, este último hallazgo está demasiado próximo del centro de la galaxia. Quizá sea posible detectar los efectos relativistas, pero las medidas serán muy inciertas.

Gallo y la polémica del SIDA

Robert C. Gallo, director del laboratorio de biología de células tumorales en el Instituto Nacional del Cáncer (INC), de los Estados Unidos, está atravesando momentos difíciles. Preeminente en el campo de la retrovirología humana, ha publicado unos 900 artículos y ha recibido los principales premios en biomedicina, salvo el Nobel.

Las suspicacias en torno a su conducta se remontan al mes de septiembre de 1983, cuando el virólogo francés Luc Montagnier, del Instituto Pasteur de París, envió a Gallo una muestra de un aislado francés del virus del SIDA, llamado LAV. Siete meses después, Gallo convocó una conferencia de prensa en la que anunció que había identificado un virus, que denominó HTLV-III, causante del SIDA. Afirmó también que había desarrollado una prueba sanguínea basada en él. En 1985 Francia demandó a los Estados Unidos por quebrantamiento de contrato, sobre la base de que Gallo había utilizado la muestra vírica francesa para crear la prueba.

Los dos gobiernos zanjaron la cuestión en 1987, acordando que Gallo y Montagnier deberían compartir el mismo crédito. Estos publicaron luego el relato oficial de sus trabajos sobre el SIDA en *Nature* y escribieron un artículo en estas páginas [véase "El SIDA en 1988", de Robert C. Gallo y Luc Montagnier; INVESTIGACIÓN Y CIENCIA, diciembre de 1988].

Las aguas se calmaron hasta que el *Chicago Tribune* publicó un artículo de John Crewdson en noviembre de 1989, en el que se afirmaba de nuevo

que Gallo había reclamado el crédito de mala fe. Este artículo movió al congresista John D. Dingell a solicitar que el Instituto Nacional de la Salud (INS) examinara el caso. En febrero de 1990, el INS (o NIH, por sus iniciales en inglés) abrió una investigación sobre el papel del laboratorio de Gallo en el aislamiento de las muestras del virus del SIDA, el crecimiento del virus en cultivos celulares y la comprobación de que era la causa principal de la enfermedad.

Después de que el grupo de expertos hubiera tomado declaración durante horas y hubiera examinado los registros del laboratorio de Gallo, William F. Raub, director en funciones del INS, anunció que el grupo había determinado que Gallo hizo "un número sustancial" de detecciones y aislamientos del virus del SIDA "a partir de varios y distintos orígenes" durante el período crucial en que él y sus colaboradores estuvieron aislando y haciendo crecer el HTLV-IIIB. El hallazgo, dice Raub, "elimina el motivo principal" por el que Gallo o uno de sus colegas pudiera haber robado el virus francés.

Sin embargo, al mismo tiempo, Raub dijo que en un informe de investigación clave, hecho por Mikulas Popovic y otros del grupo de Gallo, había "una diferencia aparentemente significativa entre el trabajo tal como se describía y lo realizado". Raub emprendió una investigación oficial que se centró en los datos que faltaban y en algunas afirmaciones sin base del artículo.

A nadie le gusta esperar la sentencia de un jurado. Gallo, molesto, quiere tener la última palabra. "Estoy seguro de que nadie de mi laboratorio cometió ninguna fechoría. Es evidente que Mika [Popovic] no toma sus notas adecuadamente, pero hizo una gran contribución: descubrió cómo desarrollar el virus. No entiendo que nadie pueda interpretar que estas cosas constituyan un crimen. ¿Qué finalidad tendría ese crimen?"

La personalidad mordaz y veleidosa de Gallo es una de las realidades de la ciencia biológica. Provoca a la gente, dicen sus colegas, con su propensión al sarcasmo y su actitud a veces irónica hacia el trabajo de otros científicos, incluyendo el grupo de Montagnier. Pero todos están de acuerdo en que Gallo es un pensador riguroso que hace esfuerzos denodados para dominar intelectualmente.



5. ROBERT C. GALLO, centro de la polémica.

Gallo nació en Waterbury (Connecticut) en 1937. La ciencia no le atraía, dice, aunque estaba acostumbrado a observar a su padre, un metalúrgico autodidacta, enfrascado en revistas técnicas. A los 12 años, vio a su hermana menor, Judith, enfermar de leucemia. Antes de que muriera un año después, el padre de Gallo la llevó al Hospital Infantil de Boston, donde fue uno de los primeros pacientes en recibir quimioterapia.

Gallo admite que la tragedia, que desoló a su padre, pudo haber influido en su carrera. Padre e hijo acabaron haciéndose amigos íntimos de Marcus Cox, el patólogo que diagnosticó la enfermedad de Judith. Gallo describe a Cox como un escéptico alegre y admite su poderosa influencia. "Yo seguía sin querer ser médico, principalmente porque mi padre quería que lo fuera", dice Gallo. "Si alguien quiere que haga algo, se me produce una reacción de anticuerpo. Pero nunca me apremió, dejó de insistir. No era como yo: era diplomático."

De joven, Gallo estaba más interesado en los deportes y las chicas, según él mismo cuenta, pero cuando entró en la Facultad de Providence la ciencia había cautivado su imaginación. Su viaje hacia el éxito (y la notoriedad) empezó con un intento ambicioso y desastroso de descubrir la acción de la glándula del timo en ratones: todos sus 200 ratones experimentales murieron. En la Facultad de Medicina de Jefferson, en Philadelphia, se saltó las clases de anatomía general y de histología. "Le habían dado un nombre a todo, el de alguien que ya había trabajado en ello." En cambio, pasó gran parte de su tiempo en el laboratorio de Allan J. Erslev, centrado en la investigación de células sanguíneas. "Me gustaba pensar en las reacciones bioquímicas que acontecían en el cuerpo."

Después de un período de médico interno en la Universidad de Chicago, Gallo se incorporó al INS en 1965 para su servicio militar. La combinación de docencia, atención a los pacientes e investigación era "la cosa más noble que podías hacer". Trabajó primero en un pabellón de leucemia infantil ("el último lugar en el que hubiera deseado estar"). Pronto aceptó un puesto en un laboratorio de leucocitos y se dedicó a la investigación de la leucemia.

Gallo recuerda su palmarés académico. El y sus colegas descubrieron el factor de crecimiento de la célula *T*, la interleucina-2, en 1976, y el primer retrovirus humano, HTLV-I, en 1979. Estos descubrimientos por sí solos le hubieran valido probablemente el

premio Nobel, creen algunos investigadores, si Gallo no se hubiera visto enzarzado en la disputa sobre el virus del SIDA. Empezó a sentirse interesado por esta epidemia en 1982 y propuso que la enfermedad debíase a una variante retrovírica del HTLV-I. La idea "estimuló la búsqueda adecuada", afirma. "Resultó ser en parte cierta y en parte errónea, pero así es como funciona la ciencia."

En mayo y junio de 1983 "intensificó mucho los esfuerzos" sobre el SIDA, redirigiendo al equipo y coordinando un asalto a la enfermedad. "Que yo sepa, fuimos el primer laboratorio del mundo en cultivar regularmente muestras del SIDA", comenta. Las reuniones anuales que tenían lugar en su laboratorio se convirtieron en la meca de los investigadores del SIDA de todo el mundo. "Entiendo que mi papel es el de criticar, plantear hipótesis, estimular e intentar cuestionar la dirección del laboratorio."

El sensacionalismo de los periodistas le exaspera. "Lo peor es observar y no hacer nada acerca de las cosas que se publican que son mentiras, verdades a medias o enteras tergiversaciones. He leído repetidamente en periódicos de Europa que todo esto era por dinero."

No pueden estar más equivocados. Cuando desarrolló la prueba sanguínea, funcionarios del Departamento de Sanidad le aconsejaron que "sacara alguna patente para protegerse contra el fraude, o si no cualquiera podría afirmar que poseía una prueba sanguínea". Aunque varios de sus descubrimientos se han transformado en técnicas médicas que "suponen dinero para otras personas", Gallo sostiene que "cuando se me dijo que patentara nunca se me dijo que podría obtener ni un céntimo por ello. En realidad, podíamos obtener solamente 10.000 dólares, y yo no lo sabía".

El INS está empleando nuevas técnicas para seguir la pista a los orígenes del virus que Gallo usó para desarrollar la prueba sanguínea norteamericana. Gallo ha reconocido que su virus pudo haber surgido de una contaminación accidental a partir del LAV, de modo que no está claro que el jurado llegue a ninguna conclusión sobre culpabilidad. Gallo asevera que su nombre será exculpado. Su laboratorio prosigue cosechando descubrimientos importantes en su labor sobre el SIDA y sobre otros virus de enfermedades humanas. "De un modo u otro, las cosas quedarán claras", promete. Si ello será a satisfacción del comité de selección del Nobel queda por ver.

Ciencia y empresa

Silicio de segunda generación

¿Qué futuro le espera al silicio, el semiconductor que nos trajo el superordenador Cray y las tarjetas de felicitación con música incorporada? Los especialistas en materiales comienzan ahora a entreverle una segunda primavera en un futuro no tan lejano, cuando los superordenadores portátiles se expendan en la tienda de artículos de electrónica del barrio.

El pesimismo precedente residía en la velocidad de los electrones; éstos se mueven a través de un cristal de silicio más lentos que en el arseniuro de galio y otros compuestos obtenidos de los elementos de las columnas tercera y quinta de la tabla periódica (materiales III-V). El silicio ofrece otros inconvenientes que hacen extremadamente difícil que emita luz, lo que concede ventaja al arseniuro de galio en un mundo donde los fotones servirán de vehículo primario para transmitir y, con el tiempo, procesar la información.

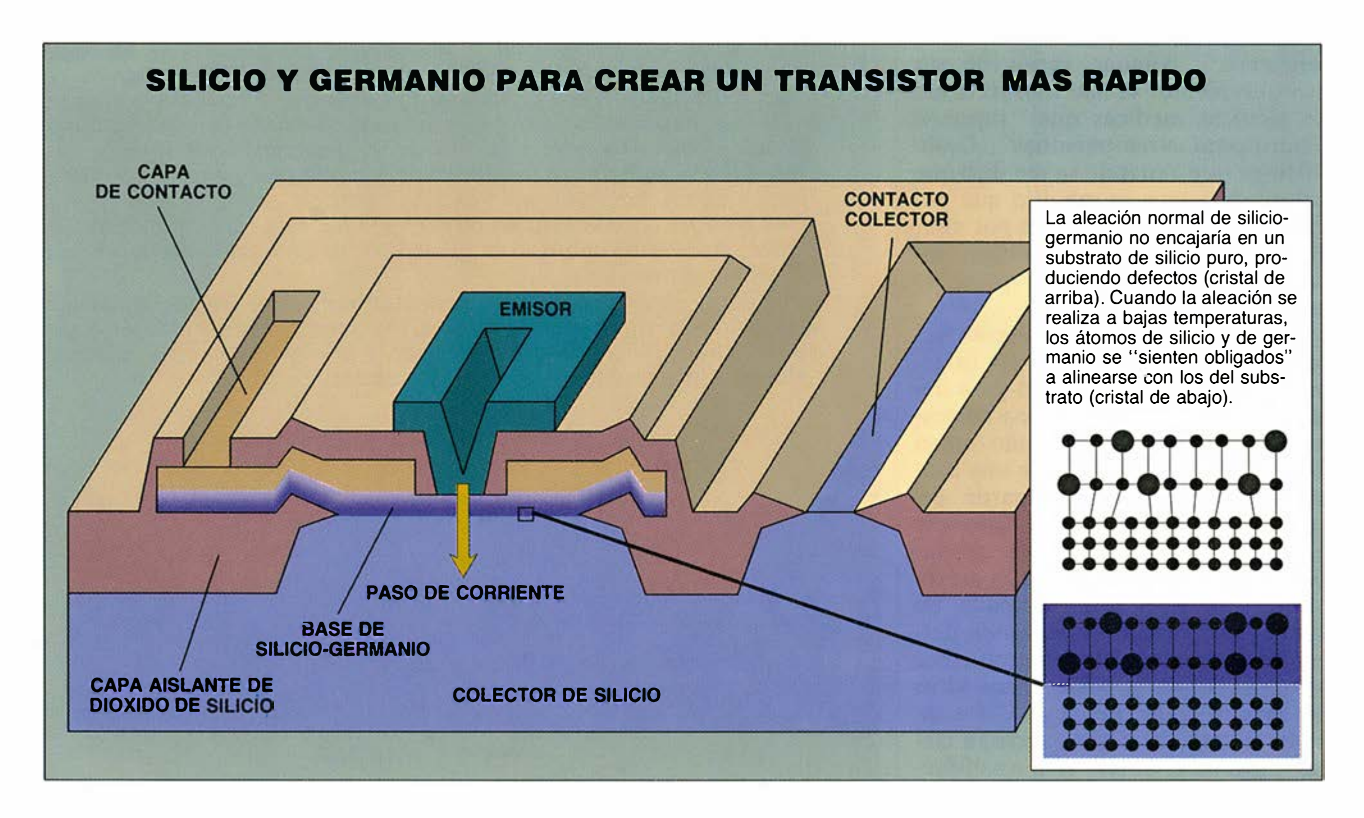
No obstante su paso de tortuga, el silicio tiene todavía probabilidad de ganar la carrera gracias a su fácil transformación en transistores y otros dispositivos electrónicos. La superioridad del silicio proviene de su capacidad de ser cortado fácilmente en grandes obleas y, cuando se calienta, de formar dióxido de silicio, aislante natural que sirve de recubrimiento cuando las configuraciones de los circuitos se graban fotolitográficamente sobre las obleas. Ni el arseniuro de galio ni ninguno de los otros materiales III-V se manejan con esa facilidad a la hora de producir dispositivos electrónicos, lo que ha impedido su empleo en gran escala pese a sus atrayentes propiedades electrónicas y ópticas.

Para continuar sacándole partido se le combina con otros materiales. La investigación en distintas empresas ha demostrado que, pese a la habitual falta de acoplamiento entre el silicio y las redes cristalinas de la mayoría de los materiales, la aleación de silicio con pequeña cantidad de germanio goza de elasticidad suficiente para que se forme sin defectos sobre un substrato de silicio puro.

La demostración más pragmática de esta idea se presentó el año pasado, cuando Gary L. Patton y colaboradores, del Centro de Investigación Thomas J. Watson de la compañía IBM, informaron haber fabricado un transistor de aleación de silicio y germanio que conmutaba a una velocidad de 75.000 millones de ciclos por segundo, casi el doble de las velocidades medidas para el silicio y cercana a las prestaciones de los transistores experimentales constituidos con materiales III-V.

La adición de germanio permitió a IBM "ajustar" el intervalo de bandas del silicio, como suele hacerse con los compuestos III-V. Llámase intervalo de bandas la cantidad de energía necesaria para elevar los electrones más externos, o de valencia, de los átomos de una red cristalina a un estado de mayor energía —la banda de conducción— en donde puedan moverse con libertad y transportar corriente. El achicamiento o el ensanchamiento del intervalo entre las bandas de valencia y conducción mejora las prestaciones de un dispositivo.

Con esta técnica, IBM fabricó un transistor bipolar, de concepción similar a los empleados en sus grandes ordenadores, usando la aleación de silicio y germanio para la base del dispositivo, es decir, la componente que permite o corta el paso de la corriente. El germanio, cuyo intervalo de bandas es menor que el del silicio, permite que pasen más electrones del emisor, sito a un lado de la base del transistor, al colector, en el otro. Con el aumento gradual de la cantidad de germanio en la base, los investigadores crearon una diferencia de potencial eléctrico que actúa como una pendiente a lo largo de la cual



los electrones se precipitan vertiginosamente.

Los avances en ingeniería de materiales abren el camino para emplear no sólo germanio, sino también una gavilla de elementos y compuestos que se pueden combinar con el silicio para crear intervalos de banda a medida. Los ingenieros de materiales pueden incluso resolver el problema, más difícil, de convertir al silicio en un ingrediente de los circuitos optoelectrónicos. A causa de su intervalo de banda, el silicio emite, por lo común, más calor que luz. Pero su ubicua presencia en electrónica lo convierte en un candidato ideal para combinar la transmisión óptica y electrónica mediante una pastilla o mediante pastillas o tarjetas de circuitos.

Convertiríamos el silicio en material primario para la optoelectrónica si lo uniésemos al arseniuro de galio u otro material III-V. Pero no es tarea fácil depositar arseniuro de galio sobre silicio. Los primeros empeños fallaron porque la alineación de los átomos individuales de arsénico y galio con los de silicio está desviada un 4 por ciento, lo que, de acuerdo con los expertos, constituye un enorme desajuste. Cuando las dos redes cristalinas tratan de alinearse, se introducen defectos en el arseniuro de galio.

Algunos investigadores continúan adelante en esa pretensión. Eli Yablonovitch y sus colaboradores, de Bell Communication Research (Bellcore), han dado, sin embargo, con una nueva manera de abordar el problema. Hace tres años usaron ácido fluorhídrico para desprender una fina capa de una oblea de arseniuro de galio y pegarla sobre silicio. Como la película se había formado sobre el arseniuro de galio original, la estructura cristalina estaba exenta de los habituales defectos. En 1989, el grupo de Bellcore anunció que había producido láseres, transistores y otros dispositivos usando este procedimiento, que permite también depositar arseniuro de galio sobre vidrio.

Hay quien pretende que sea el propio silicio el que emita luz. Pero a ese trabajo se opone un obstáculo técnico al parecer insalvable. Los compuestos III-V son materiales de intervalo de banda directo. Cuando un electrón del arseniuro de galio cae de la banda de conducción de mayor energía a la banda de valencia, el electrón se combina con un agujero, o hueco como también se le llama, cargado positivamente para emitir un fotón. En el silicio, la transición del electrón se hace mediante un intervalo de banda indirecto. En vez de expulsar un fotón, las vibraciones del cristal liberan un electrón, disipando en forma de calor su energía por la red.

Durante años los investigadores han venido ensayando otras maneras de iluminar el silicio. Entre ellas, la de agregarle compuestos de azufre, berilio u otros átomos. Esta adición no altera las propiedades conductoras del silicio, pero los átomos de la impureza conservan su capacidad de capturar un electrón viajero o un agujero próximo. El electrón puede entonces atraer un hueco. La combinación electrón-hueco, llamada excitón ligado, podría emitir un fotón.

En 1989, el Instituto de Optica de la Universidad de Rochester construyó un diodo emisor de luz usando silicio contaminado con azufre, que emite luz en la zona del infrarrojo próximo del espectro, la que se aprovecha para la comunicación por fibra óptica (1,3 micrometros). La emisión de luz más intensa se produjo, sin embargo, a temperaturas entre −223 y −193 grados Celsius. La aptitud de los átomos de impureza para ligar la pareja electrón-hueco a la temperatura ambiente es pequeña, y la red cristalina oscila a un ritmo tal que se emite poca luz.

Podría haber un camino mejor. Dennis G. Hall, profesor de Optica en la Universidad de Rochester, y Joze Bevk, de los Laboratorios Bell, se aprestan a reforzar las propiedades emisoras de luz del silicio edificando una super-red construida con capas atómicas de silicio que alternan con otras de aleación de silicio y germanio. El pequeño intervalo de banda de las capas de silicio-germanio tiende a capturar electrones y agujeros entre las "paredes", más altas, del intervalo de banda de las capas de silicio. Así aprisionados, los electrones y huecos pueden emitir un fotón. "Las super-redes tienden a mantener los electrones y agujeros en la misma vecindad", dice Hall. Ello no obstante, la optoelectrónica de silicio tiene todavía un largo camino por recorrer.

La fiebre del gen

El proyecto genoma ha despertado el viejo sueño de Eldorado. Cartografiar y con el tiempo secuenciar los 100.000 genes, uno más o uno menos, que hay en el cuerpo humano podría costar al gobierno federal de los Estados Unidos, que es el que lo puso en marcha en octubre del año pasado, un total de 3000 millones de dólares, con un promedio de 200 millones anualmente, a lo largo de los próximos 15 años. La caza del gen está en marcha en los laboratorios públicos, institutos de investigación, universidades y empresas privadas.

PROYECTO GENOMA HUMANO

MAPAS FISICOS

Centros principales

NIH: Cromosomas humanos 4, 7, 11, X
DOE: Cromosomas humanos 16, 19, 21

Trabajos en menor escala

NIH: Cromosomas humanos 1, 4, 9, 14, 17, 18, 21, Y
DOE: Cromosomas humanos 4, 5, 7, 11, 17, 22, X

SECUENCIADO

NIH: Proyectos de megabases
Secuenciación de organismos de investigación
Nuevas técnicas de secuenciación
DOE: Secuenciación de ADN
Nuevas técnicas de secuenciación

ORGANISMOS MODELO

NIH: Expresión de genes en organismos de investigación (bacterias y ratones)

INFORMATICA

NIH/DOE: Gestión de datos, análisis, accesibilidad

FUENTE: *Centro Nacional de Investigación del Genoma Humano*

La oportunidad comercial surge en cada etapa del proceso, desde la tarea inacabable de cartografiar los genes de los cromosomas hasta la no menos tediosa labor de construir copias de aquéllos para su uso en fármacos y análisis clínicos. Toda la información y métodos desarrollados con fondos federales —cuya administración corre a cargo del Departamento de Energía (DOE) y el Instituto Nacional de Salud (NIH)— deben quedar a disposición para su "comercialización agresiva".

Nadie sabe cuál es la mejor estrategia a seguir. "Falta mucho hasta poder acotar qué pertenece a mercados especializados y qué a mercados generales, qué corresponderá a la pequeña empresa y qué a las grandes", observa Mark L. Pearson, director de investigación y desarrollo de la compañía Du Pont. Por una razón obvia: nadie sabe cuáles serán las técnicas imperantes en 1995.

Desorientación que no ha impedido que las empresas comiencen ya a movilizar sus productos y técnicas. Las recién creadas Genmap y Transkaryotic Therapies (TKT) están perfilando una técnica de laboratorio encaminada a acelerar el objetivo de los cinco primeros años del proyecto: cartografiar. La técnica en cuestión descompone el genoma en fragmentos de ADN que los especialistas analizarán de manera sistemática. Una

variación en el tamaño de los fragmentos entre individuos enfermos y sanos puede indicar que el segmento contiene el gen causante del mal. Pero no todas las mutaciones deletéreas se identificarán fácilmente. Algunos genes defectuosos pueden ser del mismo tamaño que los normales y sin embargo albergar un importante cambio: la cartografía habrá de detectarlo.

"Querríamos que se nos considerase una empresa buscadora de genes", señala William T. Carroll, director de investigación de Genmap. Fundada en 1988, la empresa ofrecerá a sus colaboradores o contratantes su técnica de laboratorio para clonar segmentos de ADN de longitudes de hasta cinco millones de pares de bases. Con segmentos tan largos, los investigadores pueden cubrir mucho terreno rápidamente: por ejemplo, a lo largo del cromosoma 17; se cree que se encuentra aquí, entre unos 80 o 90 millones de pares de bases, un gen relacionado con el cáncer de mama. La firma produce también mapas reducidos de cromosomas, con sondas que realzan las secuencias repetitivas de ADN donde, a tenor de ciertos estudios, se hallarían los genes activos.

Los investigadores de la empresa TKT creen que el paso más difícil en la cartografía de los genes consiste en pasar, o "caminar", de un marcador cromosómico conocido a un gen causante de enfermedad. Se trata de uno de los métodos más lentos de laboratorio. Por eso, los investigadores tienden a buscar genes de enfermedades raras pero que muestran características innegables de herencia en vez de condiciones menos bien definidas, como las enfermedades cardíacas.

TKT ha desarrollado un método de tránsito rápido para aislar los genes deseados. Su laboratorio declina describir la técnica reservada, pero el portavoz de la empresa dice que se tarda algunas horas en vez de las habituales cuatro o seis semanas. La firma pondrá a prueba la técnica para identificar enfermedades corrientes y sin embargo genéticamente complicadas, como la hipertensión, enfermedades maníaco-depresivas y diversas formas de cáncer.

Otras empresas están siguiendo lo que se conoce por estrategia de la fiebre del oro: vender instrumentos y productos a los buscadores. Empresas recién creadas, como Genomyx, en San Francisco, y Bios en New Haven, se están enfrentando a potentes compañías tales como Du Pont, Pharmacia y Applied Biosystems, para lograr una tajada del mercado de secuenciadores automatizados de ADN. La generación existente de máquinas detecta la secuencia de genes leyendo las configuraciones formadas por los segmentos de ADN cuando proceden a través de un gel. Las mejoras se espera que lleguen fraccionadamente, con mejores enzimas y geles.

Las nuevas máquinas robóticas que fragmentan la molécula de ADN prometen automatizar el laborioso proceso manual de la secuenciación génica. Applied Biosystems se halla comprometida en una estación de trabajo robótica que manejará segmentos muy cortos. Un brazo robótico usa un sistema de jeringuilla para transvasar los productos químicos del frasco de almacenamiento al de reacción. Se procura evitar que se evapore un microlitro de líquido al ejecutar repetidos ciclos entre la temperatura ambiente y casi el punto de ebullición.

Tal vez la mayor oportunidad de negocio, y aquella con la que los biólogos están menos familiarizados, sea la creación de técnicas necesarias para enfrentarse a las enormes cantidades de datos producidos por el proyecto. Cómo manejar la información y hacerla electrónicamente accesible al usuario —la llamada informática del proyecto genoma— es el dominio de una fuerza de choque DOE/NIH dirigida por el biofísico Dieter Söll, de Yale.

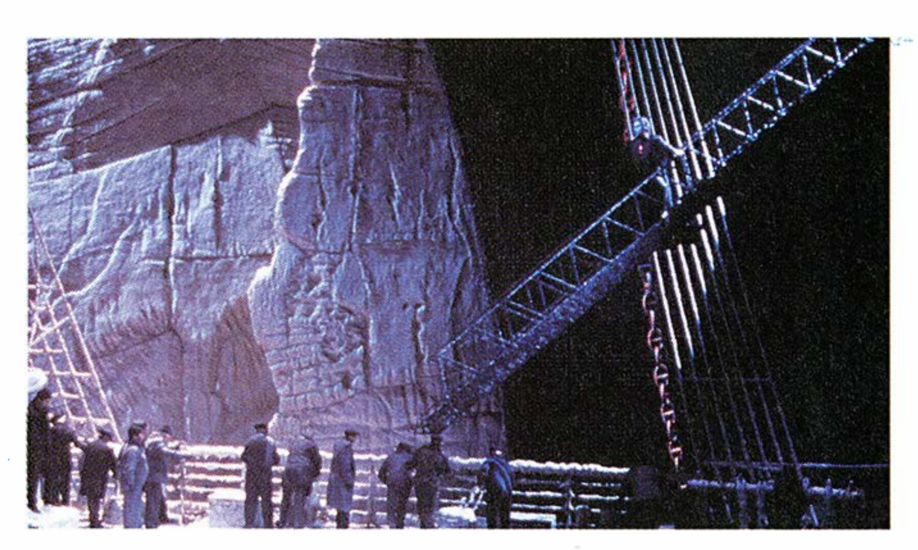
Con independencia de la técnica que una empresa esté gestando, invertir en este área es de lo más arriesgado que se pueda pensar, se dice en el sector. El riesgo no es sólo asunto de competencia y decisión rápida; habrá que contar con el aislamiento de genes que den dinero. Pero aquí se entra en el terreno de la bioética: ¿se puede o no patentar los genes?

Matemática en la escuela de teatro

Los alumnos de la Escuela de Arte Dramático de Yale aprenderán a usar una técnica matemática que sirvió para construir bombarderos de bajo perfil. Pero ellos acudirán a Broadway y a los estudios cinematográficos para crear reptiles robot de diez metros de altura y preparar superestructuras que pesan 25 toneladas cuando se cargan con luces y equipo.

Con la introducción de un curso de análisis por elementos finitos (AEF) se pretende ir más allá de las prácticas de montaje que han prevalecido desde que los tramoyistas construían chapuceramente el paraíso shakespeariano: se clavaba una tabla; si se rompía, se la sustituía por otra mayor o más fuerte. Se ha necesitado un examen más cuidadoso de la construcción de escenarios porque los decorados de Broadway, las extravagancias de las grandes salas de fiesta, la ópera e incluso el teatro local tienen ahora que estar recargados de efectos especiales y decorados laberínticos para atraer al auditorio, cansado de películas y televisión.

Se acude al análisis por elementos finitos para examinar los esfuerzos en ciertos puntos de automóviles, reactores e incluso rodilleras para los jugadores. En la construcción de escenarios, la técnica emplea matrices de ecuaciones simultáneas para deducir flexiones y esfuerzos cuando una carga —por ejemplo, una cantante de ópera de 115 kilos— se aplica a una serie de elementos finitos: las piezas



1. EL ANALISIS POR ELEMENTOS FINITOS ayudó a proyectar el decorado para El holandés errante *en la Metropolitan Opera House neoyorquina.*

de tubo, viguetas y otros componentes que forman el andamiaje de la plataforma en voladizo sobre la cual ha de estar la diva.

Sin embargo, no se exigirá a los futuros proyectistas técnicos que luchen con la resolución de ecuaciones para aprender la técnica. Algor, un programa para ordenador, hará los cálculos. El programa simplifica el proceso pidiendo al usuario que esboce un modelo tridimensional del escenario e introduzca unas pocas variables: clase de material, grueso y esfuerzo tensor entre otras. Resuelve entonces las ecuaciones y proporciona un mapa de isolíneas que indica en rojo las zonas de máximo esfuerzo.

La Metropolitan Opera de Nueva York es uno de los primeros beneficiarios de la reforma de Yale. Trabajando sobre un decorado para *El holandés errante*, Geoffrey Webb, proyectista técnico, tuvo que modificar las especificaciones de una escalera de tubo de acero de quince metros que descendía del buque flotante del holandés. El programa mostró que los escalones podían haberse doblado ligeramente hacia arriba y hacia abajo cuando el cantante hubiera avanzado sobre ella.

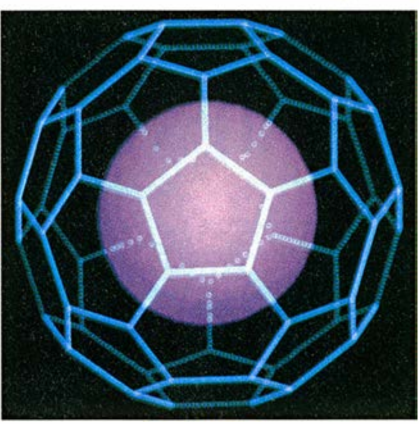
Webb busca nuevos programas en la caja de herramientas del ingeniero proyectista. Los métodos informáticos empleados para proyectar un bombardero invisible pueden ser visiblemente empleados por quienes trabajan entre bambalinas.

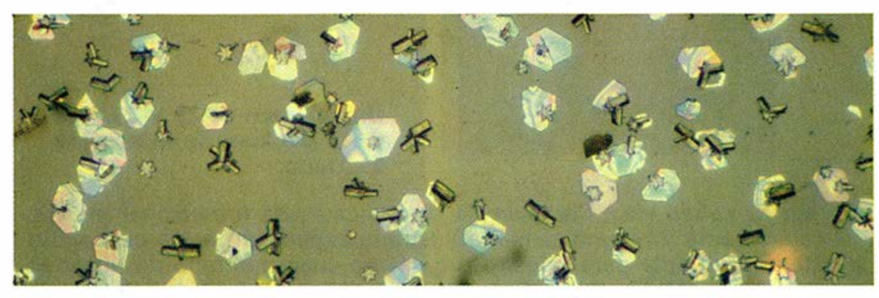
Tercera forma del carbono

Investigadores de Arizona y Alemania han dado nuevo lustre al hollín. Han aislado, a partir del mismo, ciertas cantidades de un microagregado de 60 átomos de carbono llamado buckminsterfullereno, en honor del proyectista de la cúpula geodésica. La substancia aislada, que es un cristal brillante de color pardo, constituye la tercera forma del carbono (las otras son el diamante y el grafito) y podría abrir nuevos horizontes en síntesis química. "Se ha desatado la imaginación sobre las posibilidades de uso", dice Richard E. Smalley, de la Universidad Rice, que ayudó a descubrir el microagregado y determinó su estructura. "Es un nuevo material de partida para compuestos orgánicos, semejante al descubrimiento del benceno en 1825."

La producción de cantidades del orden del gramo de buckminsterfullereno (abreviadamente, "buckeybolas") tiene dos consecuencias inmediatas. Primera, permite a los quími-







2. *LA FORMA DE BALON DE FUTBOL del buckminsterfullereno fue originariamente modelada en papel* (arriba, a la izquierda) *por Richard E. Smalley, de la Universidad de Rice. Ahora quiere "empaquetar" iones dentro de la caja de carbono-60* (arriba, a la derecha). *Los cristales de esta materia fueron recientemente aislados* (abajo) *por Donald R. Huffman, de la Universidad de Arizona, y Wolfgang Krätschmer, del Instituto Max Planck de Física Nuclear. Krätschmer suministró la microfotografía.*

cos echar su primer vistazo a una masa apreciable de una especie de carbono que se había descubierto, en forma molecular, sólo cinco años antes. Segunda, las técnicas de química de grandes masas, tales como la cristalografía, han demostrado sin lugar a dudas que los constituyentes de C-60 adoptan la forma de un balón de fútbol —semejando las cúpulas de Fuller— barruntada en 1985 por Smalley, sus colegas de Rice y Harold W. Kroto, de la Universidad de Sussex.

La técnica de producción en masa se descubrió por accidente, cuando los investigadores buscaban aglomerados todavía más pequeños. "Estábamos produciendo cantidades del orden del miligramo en 1983, pero no sabíamos lo que era; después del trabajo de Rice comenzamos a pensar si no se trataba de carbono 60", recuerda Donald R. Huffman, de la Universidad de Arizona en Tucson, que trabajaba con Wolfgang Krätschmer, del Instituto Max Planck de Física Nuclear de Heidelberg.

Huffman y Krätschmer lograron su síntesis haciendo saltar un arco entre electrodos de grafito en el seno de helio mantenido a la presión de una séptima parte de atmósfera. Recogieron el hollín que se formó en el recipiente, lo disolvieron en benceno para eliminar especies de carbono indeseadas y les quedó una mezcla casi pura de C-60 y C-70, un pariente en forma de huevo. La solución, rojiza, precipita en cristales pardos.

Si se elimina la impureza C-70, los cristales presentan un color mostaza. La teoría predecía que el C-60 sería amarillo; "la búsqueda del frasquito amarillo" duraba ya cinco años, aun cuando la "buckeybola" aparecía en los anales de química como un extraño pico en el espectrógrafo de masas. Los investigadores llegaron a la conclusión de que el pico del C-60 reflejaba la notable estabilidad del tamaño del microagregado. Se discutió sobre las posibles estructuras que podrían explicar dicha estabilidad, considerando tanto modelos abiertos, en forma de lámina, como modelos cerrados, en jaula.

Después de toscos intentos de modelado usando palillos de dientes, plastilina y decepcionantes programas de construcción de gráficos por ordenador, Smalley halló el método analítico decisivo. Unió mediante cinta adhesiva polígonos de papel en su cocina durante la madrugada de un día de agosto en 1985. Ante la obstinación de los hexágonos en su incapacidad de cerrar el recinto, Smalley se tomó un descanso para beber una cerveza. Reanudado el trabajo, halló que 12 pentágonos y 20 hexágonos creaban una bola si no se dejaba que los pentágonos compartieran un lado y si sus vértices apuntaban uno al otro (para conservar la simetría). Había tropezado con el icosaedro truncado, llamado así porque se puede construir cercenando los vértices del icosaedro, último de los "poliedros perfectos" de Platón.

La forma es perfecta. "Sesenta es el máximo número de rotaciones propias en el grupo icosaédrico", dice Smalley. "Este, a su vez, es el máximo grupo puntual, o máximo grupo en el que las operaciones de simetría, giros, reflexiones, etcétera, dejan un punto fijo. Esto convierte al C-60 en la molécula más simétrica posible."

Esta molécula, la más redonda entre las redondas, podría tener una amplia gama de aplicaciones comerciales. Podría servir como núcleo de un magnífico lubricante, predice Smalley, supuesto que sus 60 enlaces pendientes se fijen primero, quizá añadiendo 60 átomos de flúor. (Cada átomo de carbono tiene cuatro valencias, de modo que puede formar un doble enlace con uno de sus tres vecinos o combinarse con otro átomo exterior.) La simetría de la buckeybola deja libertad a los electrones para moverse por la red, que funciona a modo de una cúpula geodésica: el esfuerzo impuesto sobre las uniones rectas se reparte por toda la jaula. En un edificio, esto garantiza un techo sólido y ligero; en un microagregado de carbono, favorece la fácil cesión y adquisición de electrones, reacción reversible que Smalley cree puede crear una clase completamente nueva de baterías recargables.

Lo más emocionante de todo se cifra en la posibilidad de "empaquetar" iones metálicos en el interior de la cavidad central de la caja de carbono. Esto permitiría encapsular material reactivo o radiactivo para su almacenamiento especial. Las buckeybolas rellenas podrían también servir de diminutas asas. No es fácil tomar un solo átomo: las trampas de iones requieren que estén ionizados; pero los átomos más pequeños deben caber dentro de la buckeybola aunque no estén ionizados.

La técnica favorecería, asimismo, el ajuste fino de las propiedades electroquímicas y ópticas de la buckeybola. "En principio, podríamos alojar en su interior 92 elementos y obtener 92 propiedades diferentes", expone Smalley, que ha apodado a los nuevos compuestos metalofullerenios.

La formación de las moléculas

En menos de una billonésima de segundo, los átomos chocan, interaccionan y dan lugar a moléculas. Gracias al láser y los haces moleculares, podemos observar los movimientos durante el cambio de una sustancia a otra

Ahmed H. Zewail

En 1872, el magnate del ferrocarril Leland Stanford apostó 25.000 dólares a que un caballo a galope tendido levantaba los cuatro cascos del suelo en algún momento de la zancada. Para demostrarlo, Stanford contrató al fotógrafo inglés Eadweard Muybridge. Tras muchos intentos, Muybridge construyó un obturador que se abría y cerraba en sólo dos milésimas de segundo, permitiéndole plasmar en la película la imagen del caballo en el aire [*véase la figura 1*]. A lo largo de los últimos cien años, todas las disciplinas científicas —desde la astrofísica hasta la zoología— han sacado partido de la fotografía de alta velocidad, revolucionando el conocimiento de los movimientos animales y mecánicos más rápidos que los que el ojo humano puede seguir.

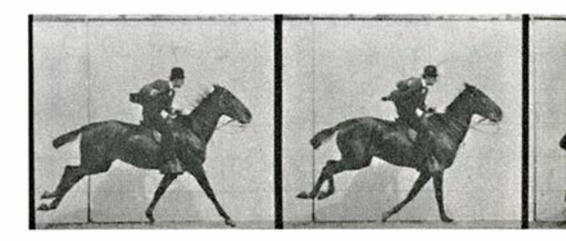
La resolución temporal —es decir, la velocidad del obturador— que se necesita para fotografiar el ultrarrápido movimiento de las moléculas escapa a cualquier escala convencional. Cuando una molécula se fragmenta o se combina con otra para formar una nueva, los enlaces químicos que ligan los átomos se rompen o se forman en menos de una billonésima de segundo (o sea, un picosegundo). Los científicos abrigaban la esperanza de poder observar los movimientos moleculares en tiempo real y presenciar el nacimiento de las moléculas: el instante en el que se decide el destino de la reacción molecular y quedan determinados los productos finales. Al igual que Muybridge, tenían que desarrollar un obturador ultrarrápido, pero que funcionara 10.000 millones de veces más deprisa que el modelo del siglo XIX.

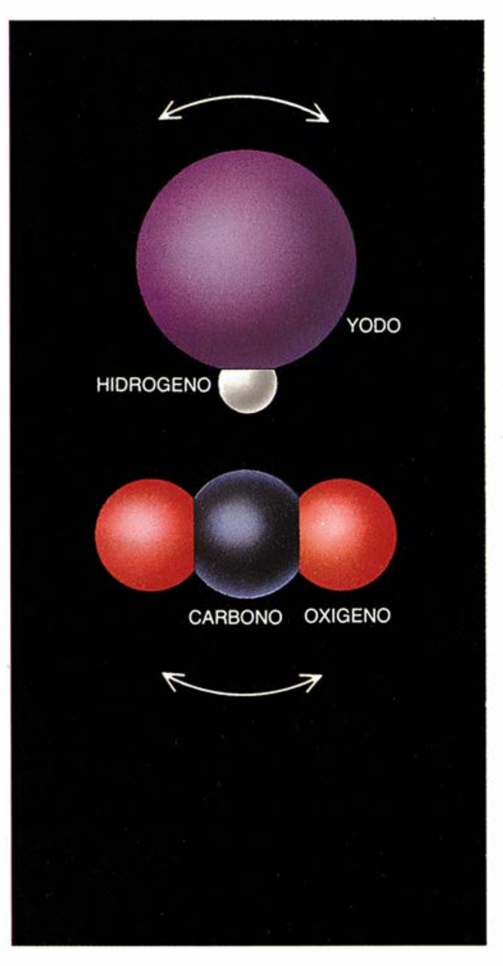
En los últimos diez años, nuestro grupo de investigación del Instituto de Tecnología de California (Caltech) ha desarrollado técnicas para observar la dinámica de las moléculas en tiempo real. Entre 1985 y 1987 refinamos nuestro sistema de láseres avanzados y haces moleculares hasta el punto de que ahora somos capaces de registrar el movimiento de las moléculas en el momento mismo de romperse y formarse los enlaces. La reacción puede observarse conforme tiene lugar, partiendo de los reactivos y pasando por los estados de transición hasta llegar finalmente a los productos: la química tal y como ocurre en la realidad.

Puesto que la existencia de los estados de transición se limita a menos de una billonésima de segundo, la resolución temporal ha de ser menor: de pocas milbillonésimas de segundo, o femtosegundos (un femtosegundo equivale a 10^{-15} segundos). El femtosegundo es una unidad de tiempo más pequeña que el «tic-tac» del reloj atómico más fino. Un femtosegundo es a un segundo lo que un segundo es a 32 millones de años. Más. Mientras que en un segundo la luz recorre casi 300.000 kilómetros (aproximadamente la distancia que separa la Tierra de la Luna), en un femtosegundo la luz sólo recorre 0,3 micras (algo así como el diámetro de la bacteria más pequeña).

Los alquimistas de la Antigüedad —de Egipto, Grecia, Arabia, China— desconocían la importancia de-

AHMED H. ZEWAIL es profesor de la cátedra Linus Pauling de química física del Instituto de Tecnología de California. Promotor de la femtoquímica, emplea láseres ultrarrápidos para investigar el movimiento de las moléculas, en tiempo real, durante las reacciones químicas. Comenzó sus estudios en la Universidad egipcia de Alejandría y se doctoró por la de Pensylvania. Dedica este artículo "a la memoria de mi querido amigo Dick Bernstein, un gigante en el campo de la ciencia de haces moleculares".





1. LA RESOLUCION TEMPORAL de la fotografía de alta velocidad se ha multiplicado por 10.000 millones a medida que la técnica ha ido evolucionando, desde registrar el movimiento de los animales hasta filmar el nacimiento de las moléculas. En un segundo, un caballo galopa 10 metros, como muestran las fotografías tomadas por Eadweard Muybridge en 1887 *(arriba)*. En cinco billonésimas de segundo, el yoduro de hidrógeno colisiona con el dióxido de carbono, creando monóxido de carbono, hidróxido y yodo. La ilustración *(abajo)* se basa en observaciones cuantitativas realizadas en el Caltech.

cisiva de la escala temporal en el cambio de una sustancia a otra, aunque conocieron el arte de la transmutación. Hasta este siglo, los químicos carecían de los medios necesarios para adentrarse en la ciencia de la dinámica de las reacciones moleculares.

A nivel molecular, una reacción comienza cuando dos moléculas (llamémoslas *A* y *B*) se acercan tanto que empiezan a interaccionar. Conforme las moléculas se aproximan cada vez más, aparecen otras especies moleculares —que no son ni *A* ni *B*— hasta que se crea una nueva molécula estable, *C*.

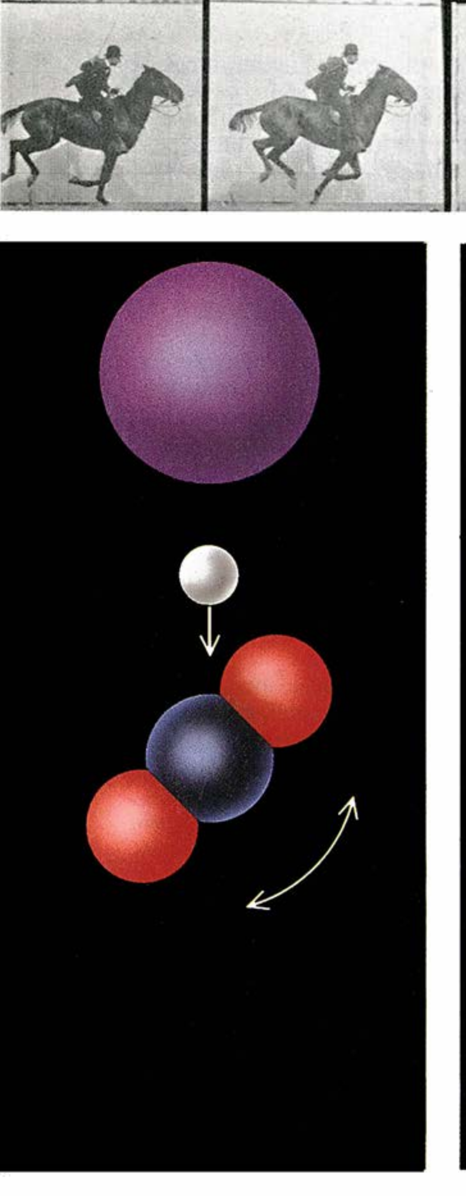
Una de estas especies transitorias constituye la etapa desde la cual la reacción procede irreversiblemente hacia los productos. A ese estado de transición, como se denomina técnicamente dicha etapa, lo representaremos aquí por $A \cdots B$. Por tanto,

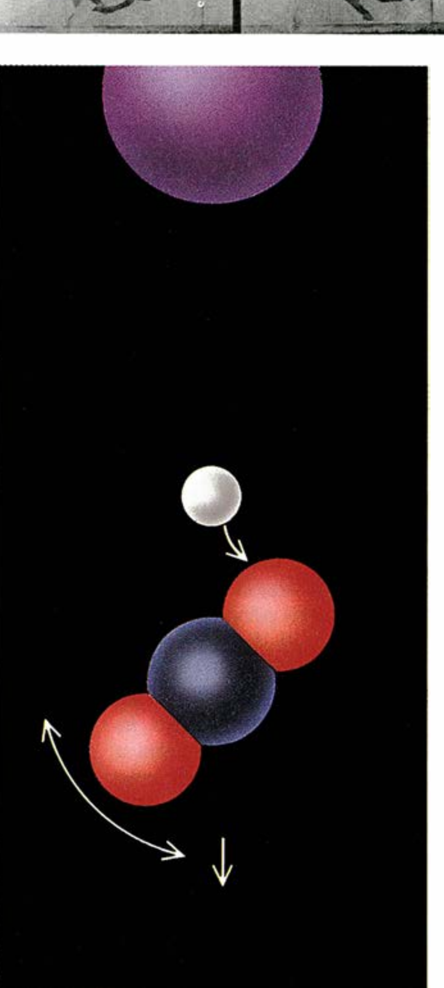
$$A + B \rightarrow A \cdots B \rightarrow C$$

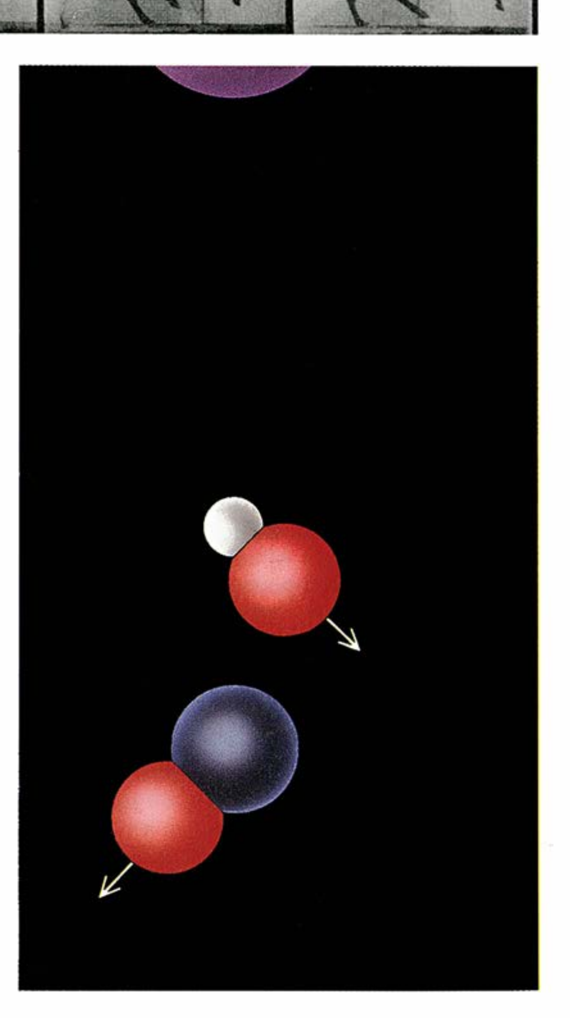
El proceso inverso es también posible: si la molécula *C* se excita, da lugar a las moléculas *A* y *B*, pasando por un estado de transición.

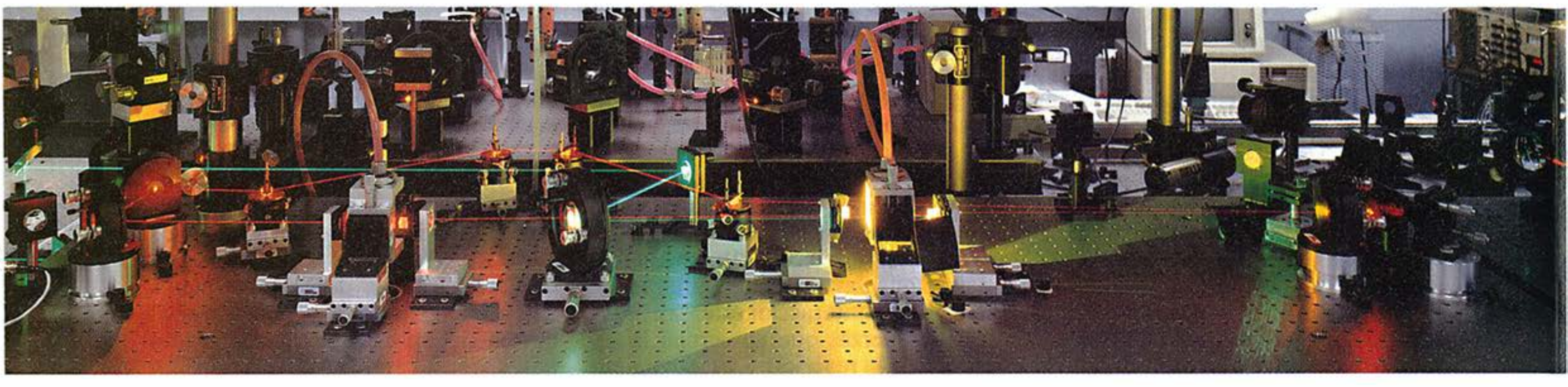
En cualquier reacción química, los movimientos de los electrones y los núcleos de los átomos determinan la forma en que las moléculas interaccionan; dichas interacciones crean, a su vez, las fuerzas que gobiernan la dinámica de la reacción. Los movimientos moleculares quedan descritos con rigor por las leyes de la mecánica cuántica. En muchos casos, la mecánica newtoniana, o clásica, puede proporcionar de manera sencilla la transferencia de energía y momento en el curso de la reacción, con la misma limpieza con que describe el choque entre dos objetos. Si los investigadores pudieran determinar cómo cambian los movimientos moleculares durante la fase de transición crítica, estaríamos en condiciones de entender cómo se forman los nuevos enlaces químicos y desaparecen los antiguos.

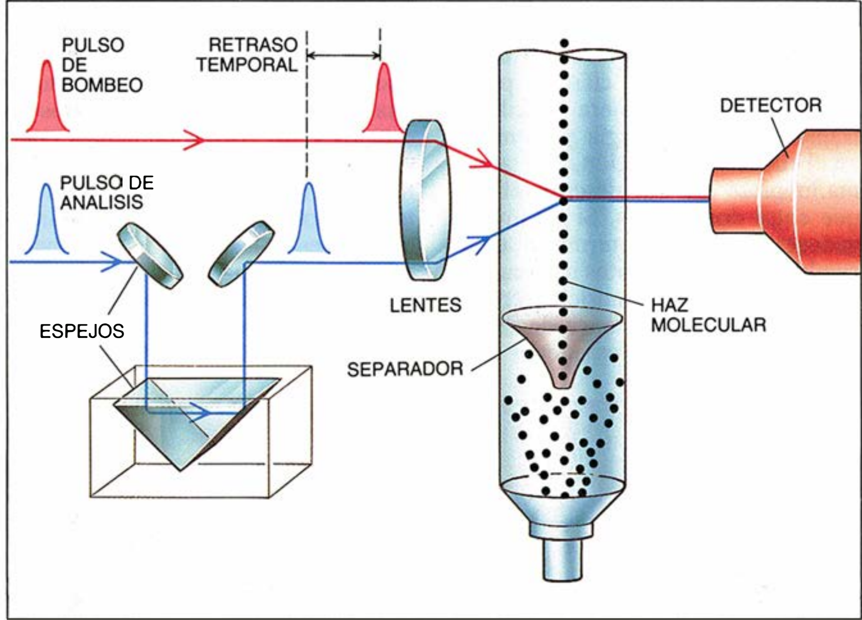
En la práctica, los químicos no siguen la pista de todos los movimientos posibles de cada uno de los núcleos y electrones del sistema molecular, pues han descubierto métodos más ingeniosos de representar el estado de un sistema. Cada estado posee cierta cantidad de energía. La energía potencial que tiene una molécula cuando los átomos se encuentran separados determinada distancia puede representarse por un punto en una gráfica. Al representar todos los estados y sus correspondientes energías potenciales se obtiene una superficie con montañas y valles. Los sistemas moleculares irán espontáneamente desde los estados de elevada energía (montañas) hacia los de energía más baja (valles), pero necesitarán energía para subir de los valles a las montañas. El fondo de un valle in-











2. FOTOGRAFIA de femtosegundo: captura la dinámica de las reacciones moleculares en tiempo real. El sistema láser *(arriba)* genera un pulso de bombeo (excitación) y otro de análisis. Tal y como se muestra en el esquema dibujado *(izquierda)*, el pulso de excitación y el de análisis salen del sistema láser al mismo tiempo. El de análisis es desviado para que entre ambos pulsos medie un intervalo temporal de pocos femtosegundos. Cuando el pulso de bombeo incide sobre una molécula, desencadena una reacción química. El pulso de análisis incide sobre la molécula pocos femtosegundos más tarde. La molécula emite entonces un espectro luminoso, en virtud de cuyo análisis el químico puede determinar la dinámica de los átomos en tiempo real.

dica un estado estable; las pendientes que rodean al valle representan la región de estados de transición, las diferentes configuraciones de los átomos cuando están próximos. Los químicos definen técnicamente un estado de transición como un punto de silla de montar en una superficie de energías potenciales.

Si un sistema molecular se compone de dos átomos y si la energía potencial del sistema no depende sino de la distancia que los separa, la superficie de energías potenciales se reduce a una curva. Si al sistema lo gobiernan la atracción de los átomos con cargas de signo opuesto a "largas" distancias y la repulsión entre los núcleos atómicos a distancias "cortas", la curva presentará un único pozo a distancias "intermedias", dando a entender que el sistema allí es estable.

La superficie de energía potencial que corresponde a un sistema molecular compuesto de muchos átomos capaces de moverse de múltiples formas puede llegar a ser complicada y pluridimensional. De todas maneras, la superficie es muy útil para describir todos los posibles caminos que pueden seguir las moléculas a partir de los reactivos (valles), pasando por los estados de transición (montañas) hasta llegar a los productos (valles).

Se puede estimar cuánto tiempo tardan los sistemas moleculares en atravesar las montañas —o sea, la duración del estado de transición— aplicando los conceptos de la mecánica newtoniana. Los átomos y moléculas pesados emergen de las reacciones químicas a velocidades típicas de 1000 metros por segundo. Si las moléculas A y B están lo bastante próximas como para hallarse en un estado de transición, entonces la distancia entre ellas es del orden de la de un enlace molecular, aproximadamente 0,1 nanómetros (o sea, 10^{-10} metros). El cociente entre esta distancia y la velocidad de retroceso nos proporciona el tiempo durante el cual A y B forman el estado de transición: 100 femtosegundos. En la más "simple" de las reacciones químicas —un átomo ligero de hidrógeno que se aproxima a una molécula de hidrógeno— la vida del estado de transición es muy corta, de unos 10 femtosegundos. Así, pues, las escalas de tiempo de los estados de transición oscilan entre los 10 y los 100 femtosegundos.

Durante más de 100 años los químicos han estudiado los mecanismos de las reacciones y la cinética molecular con el ánimo de comprender la reactividad. En el siglo pasado, el científico sueco Svante Arrhenius dio un salto importante en la descripción de la dinámica macroscópica de la reacción química. Determinó la influencia de la temperatura en la velocidad de reacción. Su famosa ecuación exponencial indica que la velocidad de reacción aumenta con el calor cedido al sistema.

Pero esas ecuaciones, que describen las propiedades dinámicas macroscópicas, sólo proporcionan información de la dinámica molecular microscópica si se admite cierto número de hipótesis. No obstante, se pueden deducir muchos aspectos de la dinámica molecular a partir de observaciones macroscópicas, de modo que la ecuación de Arrhenius conserva hoy en día su importancia.

Hace unos 40 años, se introdujeron técnicas para estudiar los intermedios químicos y la cinética química. Ronald G. W. Norrish y George Porter, de la Universidad de Cambridge, y Manfred Eigen, del Instituto Max Planck de Química Física, lograron resolver fenómenos químicos que duraban menos de una milésima de segundo. Esta escala de tiempo era ideal para los intermedios, pero demasiado larga para los estados de transición.

Para analizar la dinámica del choque de las moléculas, los investigadores desarrollaron, en los años sesenta, los nuevos métodos de las técnicas de haces moleculares. Liberando moléculas en un vacío y colimán-

dolas con un separador (*"skimmer"*), los químicos crearon haces de moléculas que no chocaban entre sí. Cuando —pongamos por caso— un haz de moléculas de *A* se cruza con otro de *B*, las moléculas de *A* colisionan con las de *B* formando un producto, *C*. La colisión completa ($A + B \rightarrow C$) se denomina reacción bimolecular. Los investigadores detectaban entonces la molécula *C*, que se creaba en diferentes cantidades y direcciones en función de la naturaleza de las fuerzas entre *A* y *B*. Conociendo las condiciones de los reactivos antes de la colisión y las propiedades de los productos tras la colisión podían deducir la dinámica del proceso.

A la inversa, un haz de moléculas de *C* puede ser bombardeado con luz láser, que excita las moléculas y rompe sus enlaces para producir los fragmentos *A* y *B*. Esta reacción de disociación ($C \rightarrow A + B$) puede analizarse midiendo las propiedades de los fragmentos. Desde que se desarrollaron los haces moleculares y otras técnicas que resuelven el estado cuántico de los productos, los químicos han aumentado bastante su conocimiento de la dinámica elemental de las reacciones. Sirva como testimonio de la importancia de la ciencia de haces moleculares y la dinámica de las reacciones la concesión del premio Nobel de química de 1986 y del premio Welch de 1988 a sendas aportaciones en estos campos.

En los experimentos de haces moleculares, la secuencia de sucesos de una reacción química, entre ellos los estados de transición, no puede observarse en tiempo real. El químico Sture Forsén, de la Universidad de Lund, dio con una clarificadora analogía que ilustra la importancia que tiene comprender los estados transitorios en la dinámica. Comparaba a la comunidad científica con los espectadores de una versión drásticamente recortada de un drama clásico. A los asistentes se les ofrecen sólo las primeras escenas de *Hamlet,* por ejemplo, junto con el final. Escribe Forsén: "Los personajes principales se presentan, entonces baja el telón para cambiar los decorados y, cuando se levanta de nuevo, nos encontramos en el escenario con un considerable número de muertos y unos pocos supervivientes. No es precisamente una tarea fácil para el que no es experto descubrir la trama de lo sucedido".

En el intento de analizar los estados de transición más directamente de lo que permiten los experimentos convencionales de haces moleculares, los químicos han adoptado técnicas recientes y han desarrollado métodos para registrar la absorción, emisión o dispersión de la luz por las moléculas de los estados de transición. Al igual que ocurre con los haces moleculares, estos métodos no resuelven los sucesos en tiempo real, pero ofrecen oportunidades más directas para examinar de cerca la dinámica y los estados de transición.

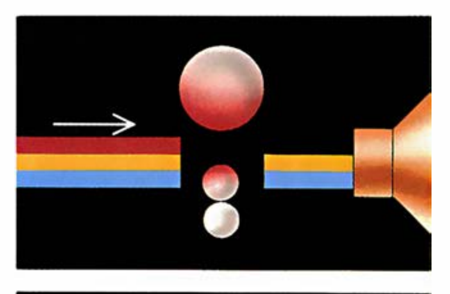
En 1979 nuestro grupo de investigación del Caltech abordó por primera vez el estudio de los fenómenos químicos ultrarrápidos usando láseres y haces moleculares. Por aquel entonces la resolución temporal era de decenas de picosegundos, una escala demasiado basta para ver el paso de las reacciones a través de los estados de transición. Nos concentramos en el desarrollo de técnicas para estudiar el flujo de energía en las moléculas antes de que tuviera lugar la reacción química. Hacia el año 1984, la resolución temporal de los experimentos con haces moleculares ya era de tres picosegundos. Este avance nos permitió medir directamente la velocidad de reacción conforme la molécula cambiaba de un estado cuántico a otro. Pero necesitábamos todavía una técnica que pudiera registrar el espectro de los estados de transición en tiempo real, con resolución de femtosegundos.

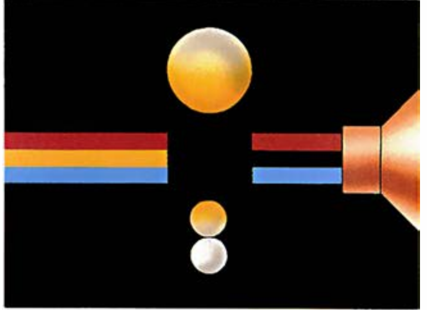
Al objeto de construir esta "cámara" molecular de femtosegundos, combinamos la técnica de haces moleculares con la de los láseres de pulsos ultracortos. Los haces moleculares nos permitían aislar la reacción en el vacío; los pulsos láser ultracortos nos brindaban la posibilidad de resolver la dinámica con la resolución temporal requerida. A lo largo de los últimos cinco años, los avances en la técnica láser conseguidos en los laboratorios AT&T Bell y en otras instituciones han permitido alcanzar la resolución temporal crítica del femtosegundo. En este momento, se pueden obtener pulsos del orden de seis femtosegundos. Con pulsos tan rápidos, se logra una "velocidad de obturador" de 10 femtosegundos.

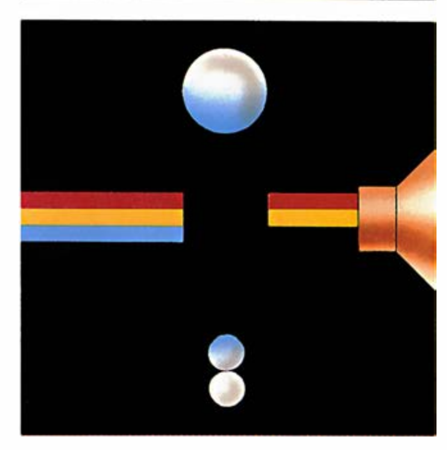
Los principios en los que se fundamenta la fotografía molecular ultrarrápida guardan cierta similitud con los aplicados por Muybridge. La clave de su trabajo era un obturador especial que exponía la película durante 0,002 segundos tan sólo. Muybridge dispuso el experimento montando doce de esas cámaras con una separación de medio metro a lo largo del recorrido del caballo. Atado a cada cámara había un hilo que atravesaba la pista y estaba conectado a un mecanismo, de manera que cuando el caballo rompía el hilo se abriría el obturador.

Con este sistema Muybridge alcanzó una resolución en cada fotografía del orden de dos centímetros, suponiendo que el caballo galopara a 10 metros por segundo. (La resolución, o definición, es sencillamente la velocidad del movimiento multiplicada por el tiempo de exposición.) La velocidad del movimiento dividida por la distancia entre las cámaras equivale al número de secuencias por segundo, 20 en este caso. El movimiento en cada secuencia queda más realzado a medida que aumenta la velocidad del obturador. La resolución del movimiento mejora conforme se reduce la distancia entre las cámaras.

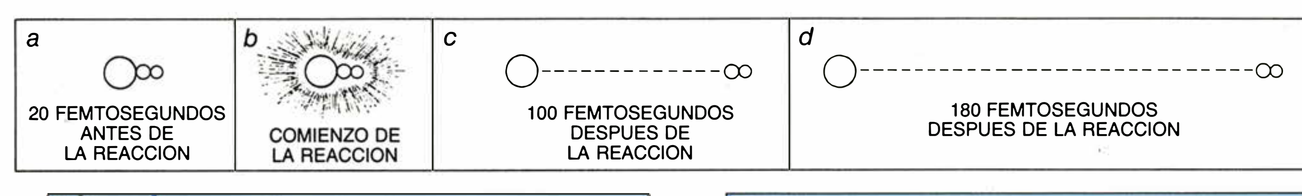
Dos aspectos de la fotografía de alta velocidad son relevantes para la fotografía molecular de femtosegundo. En primer lugar, ambos tipos de

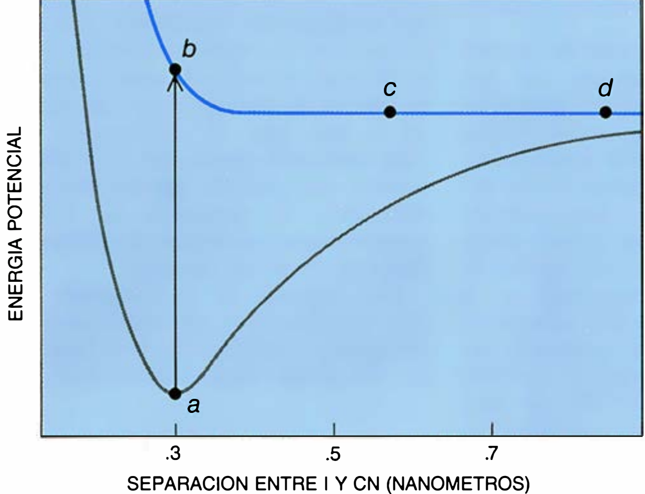


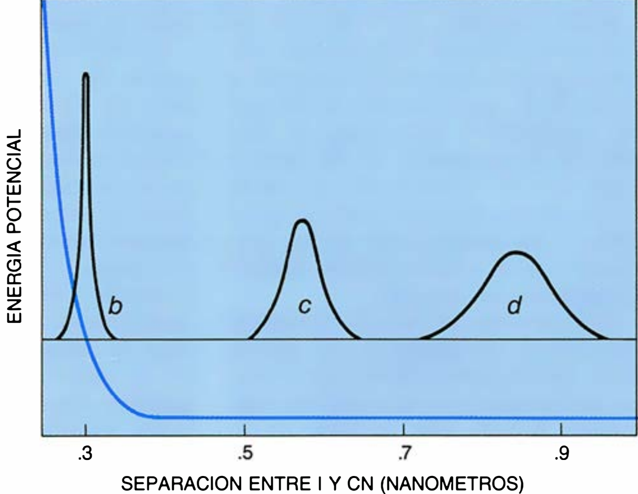




3. ESTADO de los sistemas moleculares, identificado gracias a la luz que absorben las moléculas. Cuando los átomos de una molécula están relativamente próximos, suelen absorber luz de longitudes de onda largas (como la roja). Cuando los átomos están más separados, tienden a absorber luz de longitudes de onda más cortas (azul, por ejemplo). Los cambios del espectro vienen a ser las huellas dactilares de los átomos en movimiento. (Recreación gráfica de Hank Iken.)







4. DINAMICA DEL FEMTOSEGUNDO del cianuro de yodo (ICN), descrita en función de las curvas de energía potencial. Según la mecánica clásica, cuando las moléculas de ICN de un estado de baja energía *(a)* interaccionan con un haz de láser, la luz excita el ICN hasta un estado de alta energía *(b)*. La molécula se disocia entonces en yodo y cianuro, y la distancia entre el yodo y el cianuro aumenta (*c* y *d*). El diagrama de la izquierda es una simplificación; según establece la mecánica cuántica, no puede determinarse simultáneamente el momento y la posición exacta de los átomos. Como se muestra en el diagrama de la derecha, lo más probable es que el yodo se separe del cianuro unos 0,85 nanómetros al cabo de 180 femtosegundos *(d)*. Pero en ese instante existe una pequeña probabilidad de que se hallen separados por 0,7 nm o incluso por 1,0 nm.

fotografías descomponen el movimiento continuo en una serie de instantáneas o escenas. Así se puede decelerar a voluntad un movimiento rápido para que el ojo pueda seguirlo. En segundo lugar, ambos métodos deben producir una rápida sucesión de instantáneas para que las escenas recompuestas den la apariencia de movimiento continuo. El cambio de posición de un objeto de una instantánea a otra debe ser gradual, y son necesarios como mínimo treinta fotogramas para lograr un segundo de animación.

En la fotografía molecular de femtosegundo, la definición del fotograma y el número de fotogramas por segundo deben ajustarse para resolver los movimientos nucleares elementales y, sobre todo, los efímeros estados de transición. La definición de la imagen debe ser inferior a 0,1 nanómetros. Teniendo en cuenta que la velocidad del movimiento molecular se cifra en torno a los 1000 metros por segundo, la resolución del obturador debe ser del orden de 100 femtosegundos o menos.

La idea que subyace bajo estos experimentos de femtosegundos es bastante sencilla. Un primer pulso láser, llamado de bombeo o de excitación, incide sobre una molécula aislada, iniciando la reacción y marcando el tiempo cero de nuestro experimento. Un segundo pulso, de análisis, llega varios femtosegundos después y registra una instantánea de la reacción en ese instante concreto. Al igual que las cámaras del experimento de Muybridge, la cámara molecular de femtosegundo recoge imágenes sucesivas en diferentes momentos para obtener información de los diferentes estadios de la reacción.

Para producir los retrasos temporales entre el pulso de bombeo y el de análisis, nuestro grupo de investigación aprovecha la velocidad de la luz, grande, bien que finita. Inicialmente se sintoniza el sistema óptico de manera que ambos pulsos alcancen el aparato al mismo tiempo. Desviamos entonces el pulso de análisis para que recorra una distancia mayor que el pulso de bombeo antes de alcanzar el haz molecular. Si el pulso de análisis recorre una micra más que el pulso de bombeo, se retrasará 3,33 femtosegundos, pues la luz viaja a 300 millones de metros por segundo.

Así, pues, los pulsos separados por distancias de 1 a 100 micras resuelven el movimiento durante períodos que van de 3,33 a 333 femtosegundos. De este modo, una diferencia espacial de micras, fácilmente obtenible en el laboratorio, puede aprovecharse para conseguir retrasos temporales de femtosegundos. (Una velocidad de obturador de unos pocos femtosegundos escapa a las posibilidades de cualquier cámara basada en dispositivos mecánicos o eléctricos.)

Cuando el pulso de análisis choca contra una molécula, no transmite entonces una imagen al detector, al modo de la luz que se refleja en el objeto y recoge la cámara; antes bien, el pulso interacciona con la molécula y ésta emite un espectro de luz. El espectro varía en intensidad y color. La variación depende de la longitud de onda del pulso inicial de análisis, de los átomos que forman la molécula de partida y del estado en que se hallase la molécula cuando recibió el pulso. Las moléculas libres *A* y *B* tienen un espectro diferente del que poseen *A*···*B* en el estado de transición, y el espectro del estado de transición difiere del que exhibe el producto *C*.

Pudimos, pues, distinguir las señales espectrales de cada uno de esos estados. Determinamos la pauta temporal de las mismas porque conocíamos el tiempo transcurrido entre los impactos del láser de bombeo y el de análisis. Valiéndonos de estos principios, reconstruimos y estudiamos la imagen de reacciones elementales en tiempo real.

Cuando mi colega y amigo Richard Bernstein, de la Universidad de California en Los Angeles, oyó hablar de la cámara molecular del femtosegundo, se entusiasmó mucho con ese ingenio, e imaginamos las excitantes

posibilidades que se nos abrían con esa técnica. Brindamos con champagne en su casa de Santa Mónica para celebrar el nacimiento de la "femtoquímica", como dio en llamarla.

Desde 1985 venimos investigando la femtoquímica de las reacciones de cianuros, atmosféricas y salinas. Estos ejemplos ilustran el gran alcance del método y las posibilidades de comprender mejor la dinámica de procesos químicos muy variados.

Empezamos, hace seis años, con la disociación del cianuro de yodo:

$$ICN \rightarrow I \cdots CN \rightarrow I + CN$$

La meta era detectar el fragmento CN (cianuro) en el momento de su nacimiento. Nuestros estudiantes de doctorado y postdoctorado trabajaban con empeño para lograr esa meta, y era nuestra determinación obtener resultados a finales de 1985. Construimos un sistema láser de dos pulsos y, con la ayuda de algún equipo que pedimos prestado para comprimir los pulsos, pudimos poner en marcha el experimento. En diciembre comunicábamos que habíamos detectado moléculas de cianuro en las etapas de formación. Nos dimos cuenta de que los estados de transición del cianuro de yodo quedaban un poco más allá de la resolución temporal de nuestro aparato.

Al objeto de mejorar la resolución temporal, decidimos emprender la construcción de un sistema mejor de láseres y haces moleculares. Dispusimos nuestro experimento en una sala blanca (sin polvo), totalmente rehabilitada, que había alojado en tiempos los aparatos de rayos X de Linus Pauling cuando éste se encontraba en el Caltech. Inauguramos las instalaciones poco después del Día de Acción de Gracias de 1986. Estábamos bastante excitados ante la idea de poder observar directamente los estados de transición.

Una característica crucial de estos experimentos era la de lograr que la definición de la imagen fuera suficiente para resolver el movimiento. Conseguimos un detector tan sensible que podía analizar la molécula de cianuro cuando estaba libre o en el proceso de separación del átomo de yodo. Alcanzamos esta sensibilidad valiéndonos de láseres analizadores de diferentes colores, de acuerdo con el espectro de las moléculas. Así fuimos capaces de observar los fotogramas espectrales que cambiaban con el paso del tiempo conforme procedía la reacción, de una molécula intacta de ICN a un átomo de yodo y un fragmento molecular, cianuro. Disfrutábamos comentando los resultados hasta bien entrada la noche.

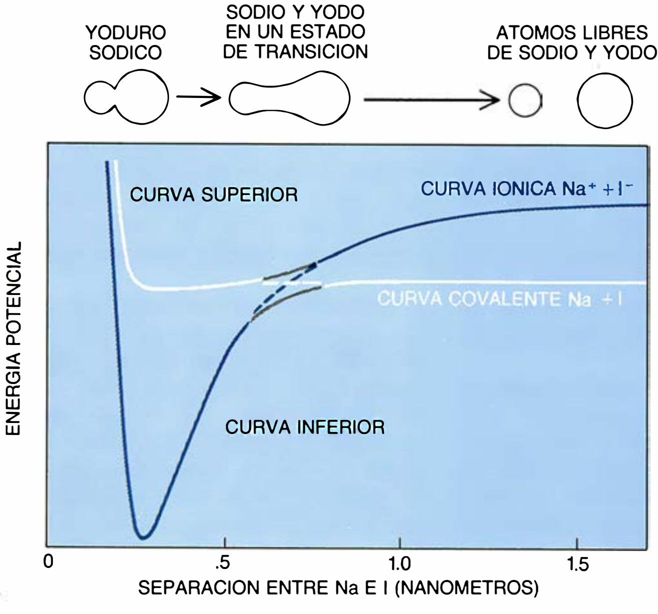
En 1987 comunicábamos el análisis de la reacción de ICN que producía yodo y cianuro, con una resolución temporal que permitía observar directamente el estado de transición $\{I \cdots CN\}$. Obtuvimos la dinámica de la reacción en tiempo real y dedujimos luego las características específicas de la superficie de energía potencial [*véase la figura 4*].

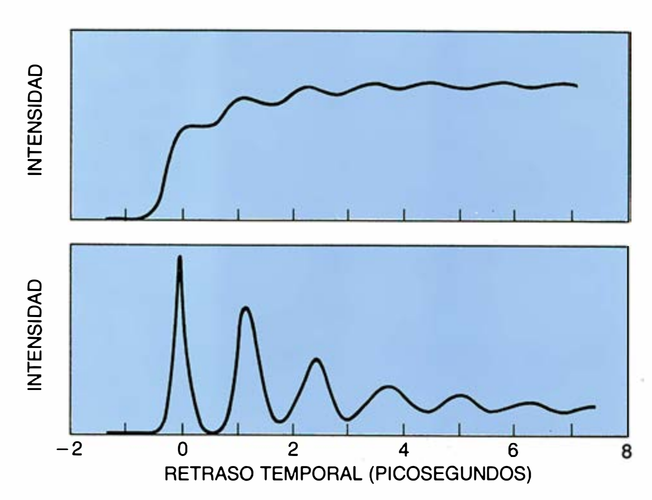
Este experimento nos permitió cronometrar la ruptura de un enlace químico elemental en una escala temporal de femtosegundos. Estos hallazgos despertaron un gran entusiasmo en la comunidad científica e inspiraron muchos otros estudios teóricos y experimentales.

En 1986, mientras estábamos aún enfrascados en la construcción de la nueva instalación de láseres de femtosegundos, Bernstein nos hizo una visita al Caltech. Planeamos entonces analizar la colisión y enlace entre dos moléculas, es decir, una reacción bimolecular. Al principio, no dábamos con la manera de detectar el inicio de esas reacciones (el tiempo cero). Aunque las reacciones bimoleculares duran menos de un picosegundo, antes las dos moléculas que reaccionan deben viajar para encontrarse, y el trayecto dura como un millón de veces más que la reacción misma.

Encontramos una solución a este problema cuando estudiamos las investigaciones de nuestros colegas de Orsay y de la Universidad de California del Sur (USC). Podíamos obviar el largo viaje y establecer inequívocamente el tiempo cero anclando los dos reactivos mediante un enlace débil de van der Waals. Por fortuna contábamos con el asesoramiento de investigaciones de la USC que habían llevado a cabo estudios (integrados, no resueltos, en el tiempo) de la reacción de bromuro de hidrógeno y dióxido de carbono.

En nuestro primer experimento en tiempo real de estos tipos de reacciones bimoleculares, combinamos yo-





5. MOVIMIENTOS DE FEMTOSEGUNDO del yoduro sódico, una sal; revelan la dinámica molecular del enlace químico que liga el sodio con el yodo. En el diagrama de energía potencial en función de la separación entre yodo y sodio, de la izquierda, la curva iónica (atractiva) interseca la curva covalente (repulsiva). Resulta, por tanto, que el yoduro sódico puede fragmentarse en yodo y sodio, actuando de manera covalente, o bien el yoduro de sodio puede existir en un estado ligado de alta energía alternando los comportamientos iónico y covalente. En el caso de que el yoduro sódico se disociase en átomos, la intensidad de parte del espectro crecerá escalonadamente, como aparece en los resultados experimentales *(arriba, a la derecha)*. Si el yoduro sódico se encontrara en el estado ligado de alta energía, entonces la intensidad de parte del espectro oscilará *(abajo, a la derecha)*.

duro de hidrógeno con dióxido de carbono para producir yodo, monóxido de carbono e hidróxido:

$$HI + CO_2 \rightarrow I + HOCO \rightarrow I + OH + CO$$

Queríamos poner de manifiesto, a nivel elemental, esta reacción atmosférica. Rompimos el enlace entre el hidrógeno y el yodo con un pulso de bombeo. Observamos entonces con pulsos de análisis que el hidrógeno atacaba el dióxido de carbono y quedaba fijado a él durante cientos de femtosegundos. Tras ello el átomo de hidrógeno despojaba al de carbono de uno de sus dos átomos de oxígeno. La molécula de hidróxido emergía finalmente cinco picosegundos después del comienzo de la reacción, como informábamos en 1987.

La técnica también nos permitió representar la dinámica del complejo transitorio de la colisión HOCO en tiempo real. Pudimos relacionar este complejo con la superficie de energía potencial que señala la aproximación del hidrógeno al dióxido de carbono. Todavía estamos examinando las muchas diferencias sutiles que se interponen entre nuestros resultados y la teoría.

Así como los físicos prefieren trabajar con el átomo de hidrógeno por su simplicidad, a los químicos físicos les gustaría comprender los sistemas moleculares elementales. La reacción química más sencilla es aquella en la que intervienen dos átomos. Me pareció que sería interesante intentar descifrar la dinámica de las moléculas de una sal, por ejemplo la de yoduro sódico:

$$NaI \rightarrow Na\cdots I \rightarrow Na + I$$

Las reacciones de los haluros de álcali fueron el prototipo más socorrido en la "edad alcalina" de los experimentos de haces moleculares, y no carecía de sentido su estudio en la "edad del femtosegundo". Además, en la Universidad de Toronto habían investigado los espectros de emisión de la molécula de yoduro sódico durante su disociación para dar los átomos de sodio y yodo.

Del yoduro sódico nos intrigaba su superficie de energía potencial, bastante interesante para la interacción entre los átomos de yodo y sodio. Si estos átomos se aproximan, se repelen entre sí, en particular a distancias cortas. Si, por otro lado, se acercan en forma de iones de cargas opuestas, se atraen mutuamente. A distancias suficientemente cortas se formará una sal compuesta por iones de sodio y yodo. Cuando el sodio y el yodo forman un enlace covalente, comparten electrones creando un estado energético estable. Cuando los átomos forman un enlace iónico, el átomo de sodio cede un electrón al átomo de yodo, originando iones cargados que se atraen.

La naturaleza, sin embargo, no obra exactamente de esta manera. En realidad los átomos se comportan como si fuesen a la vez iónicos y covalentes. Las verdaderas curvas de energía potencial son, por tanto, una mezcla de una curva iónica y otra covalente [*véase la figura 5*]. La curva iónica corta la covalente a cierta distancia. En esta zona de interacción, el sistema molecular presenta una mezcla de características iónicas y covalentes; en otras palabras, la molécula tiene cierta probabilidad de ser iónica y cierta de ser covalente.

Los teóricos han dividido las curvas de energía potencial en dos partes, una inferior y otra superior. La curva inferior, que representa los estados de baja energía, se compone de la curva iónica a cortas distancias y de la curva covalente a largas distancias. La curva superior representa la situación opuesta: es covalente a distancias cortas e iónica a distancias largas; representa, a su vez, un estado de alta energía.

El comportamiento descrito por la curva inferior es el que suele dominar cuando los átomos de sodio y yodo se acercan lentamente para que empiecen a interaccionar. En este caso, el electrón pasa del sodio al yodo creando una molécula estable de yoduro sódico. Si, por el contrario, los átomos se acercan muy rápidamente, tienen cierta probabilidad de saltar de la superficie inferior a la superior. La superficie superior no representa ya la curva covalente repulsiva (o ligeramente enlazante); posee, en cambio, un punto donde la energía es la menor posible (un mínimo). En este mínimo de potencial las moléculas serán, por tanto, bastante estables. (Con mayor rigor: las moléculas se hallan en un estado cuasienlazante.)

Confiábamos en ver la dinámica del femtosegundo del enlace del yoduro sódico a medida que se rompía, dando lugar a átomos de yodo y sodio. En particular, nos concentramos en el salto del potencial superior al inferior, o al revés. Observamos el movimiento a medida que las moléculas mudan de covalentes a iónicas y a medida que atraviesan la región de cruce. El átomo de sodio le cedía un electrón al átomo de yodo a una distancia de 0,7 nanómetros. Lo que había ocurrido es que el átomo de sodio había empleado su electrón a modo arpón, para girar en torno al átomo de yodo.

A la hora de abordar el experimento, comenzamos por enviar un pulso láser para excitar las moléculas de NaI. Cuando el sodio se separa unos 0,25 nanómetros del yodo, el enlace empieza a romperse. Se envía entonces un segundo pulso para analizar la reacción cuando ya se ha roto totalmente el enlace del yoduro de sodio y se han formado los átomos de yodo y sodio. El pulso de análisis excita el átomo de sodio, haciendo que emita luz amarilla. Analizando el movimiento desde que se rompe el enlace (tiempo cero) hasta que aparecen los átomos de sodio libres, podemos contar, en tiempo real, el número de átomos de sodio que han aparecido en la curva de potencial inferior. Y lo que es aún más importante, detectando el estado de transición cuasienlazado, Na···I, que tiene propiedades espectrales diferentes de las del sodio libre, podemos observar la transición del Na···I a átomos libres de yodo y sodio.

Al alcanzar la zona de cruce, algunas de las moléculas de yoduro de sodio saltarán a la curva inferior, y su enlace se romperá, formándose entonces átomos de yodo y sodio. Las moléculas que no saltan, sino que permanecen en la curva superior, no rompen el enlace y continúan vibrando hasta que acaban por saltar. Por tanto, los pulsos de los átomos de sodio sólo se pondrán de manifiesto tras cada viaje de vuelta en el pozo de la curva superior. Y este es exactamente el fenómeno que hemos observado. Los resultados muestran el movimiento de la molécula durante la ruptura del enlace químico y proporcionan los detalles de la energía potencial que gobierna el movimiento de los átomos de yodo y sodio.

Los laboratorios de la empresa (IBM) y de la Universidad de Friburgo están aplicando técnicas de femtosegundo para observar la dinámica de muchos tipos diferentes de reacciones en fase gaseosa. En el Caltech, seguimos analizando reacciones elementales, así como otras que dan lugar a formaciones múltiples (v.g.: $ABA \rightarrow AB + A \rightarrow A + B + A$). Esperamos también estudiar la más simple de las reacciones, la combinación de hidrógeno atómico con hidrógeno molecular.

Estamos explorando también sistemas moleculares mayores que nos permitan resolver algunas cuestiones de reactividad selectiva. Tomemos un ejemplo. Si si excita una molécula con dos enlaces iguales, ¿se rompen los dos simultánea o consecutivamente? Para adentrarnos en asunto tan general e importante, estamos estudiando la disociación del $C_2I_2F_4$ en yodo y C_2F_4. (En el C_2F_4 se forma un enlace doble entre los átomos de carbono.) A partir de los estudios en tiempo real, hemos descubierto que la reacción procede consecutivamente, a pesar de la equivalencia entre los dos enlaces entre yodo y carbono. Más notable resulta la discrepancia en la pauta temporal de ruptura de los enlaces: el primero se rompe en menos de un picosegundo, mientras que el segundo ¡tarda 100 veces más en romperse!

La capacidad de visualizar la dinámica molecular sugiere también nuevos métodos de controlar las reacciones. Hay varios esquemas teóricos para lograr el control, y en los laboratorios se están ya explorando nuevas posibilidades. Cabe esperar que pueda ajustarse con precisión el movimiento y reactividad de las moléculas. Si ello se hiciera realidad en las próximas décadas, tendríamos la química con láser a medida del cliente.

Muybridge inventó la técnica de la fotografía de alta velocidad por puro placer. La técnica ha ido mucho más allá de lo que él o Stanford pudieron imaginar. Aunque nadie puede predecir el futuro, estamos convencidos de la importancia y la belleza que encierra el conocimiento de las moléculas y sus reacciones. Queda abierta la puerta a una amplia investigación, teórica y experimental, y a hallazgos esperados. Confío en que los lectores compartan el placer y entusiasmo que hemos experimentado en el aprendizaje de los arcanos de la alquimia (la transmutación de una sustancia en otra) observando los movimientos ultrarrápidos de las moléculas.

BIBLIOGRAFIA COMPLEMENTARIA

INVESTIGATION OF ULTRAFAST PHENOMENA. Charles V. Shank en *Science*, vol. 233, n.º 4770, págs. 1276-1280; 19 de septiembre de 1986.

THE NOBEL PRIZE FOR CHEMISTRY en *Les Prix Nobel 1986*. Almquist & Wiksell International, 1987.

REAL-TIME LASER FEMTOCHEMISTRY: VIEWING THE TRANSITION FROM REAGENTS TO PRODUCTS. Ahmed H. Zewail y Richard B. Bernstein en *Chemical and Engineering News*, vol. 66, n.º 45, págs. 24-43; 7 de noviembre de 1988.

LASER FEMTOCHEMISTRY. Ahmed H. Zewail en *Science*, vol. 242, n.º 4886, págs. 1645-1653; 23 de diciembre de 1988.

EXPOSING MOLECULAR MOTIONS. Ian W. M. Smith en *Nature*, vol. 343, n.º 6260, págs. 691-692; 22 de febrero de 1990.

ULTRAFAST MOLECULAR REACTION DYNAMICS. L. Khundkar y Ahmed H. Zewail en *Annual Review of Physical Chemistry*, vol. 41, págs. 15-60; 1990

El legado de la psicología de la forma

Desde su comienzo, a principios de este siglo, la teoría de la Gestalt ha hecho importantes aportaciones al estudio de la percepción, el aprendizaje y la psicología, aportaciones que mantienen su influencia

Irvin Rock y Stephen Palmer

Como muchos movimientos científicos importantes, el de la psicología de la forma *(Gestalt)* nació de una rebelión contra la ciencia establecida de su época. En la actualidad, varias de las concepciones que a comienzos de nuestro siglo propusieron los psicólogos "guestaltistas" están ya incorporadas a la interpretación moderna de la percepción, del aprendizaje y del pensamiento. E incluso pertenecen a nuestro mismo lenguaje y a nuestra cultura. ¿Quién no ha oído alguna vez decir "en todo su conjunto", o no ha visto figuras que demuestran los principios guestálticos, como, por ejemplo, la que de pronto parece una copa pero también los perfiles de dos caras opuestas? Sin embargo, fuera del ámbito de la psicología académica, no es tan sabido en qué consistió aquel movimiento o qué se ha hecho de las ideas que le sirvieron de base.

La psicología de la *Gestalt* tuvo su origen en Alemania. Tras el ascenso del nazismo, sus fundadores –Max Wertheimer, Wolfgang Köhler y Kurt Koffka– se trasladaron a los EE.UU., donde algunos de sus discípulos siguen aún en activo. Los guestaltistas contribuyeron más al estudio de la percepción que a otras áreas de la psicología: el término alemán *Gestalt* equivale a "forma" o "figura", aunque su intención significativa quizá se traduzca mejor por "configuración". Realizaron también importantes adelantos en educación, aprendizaje, pensamiento y psicología social. Algunas de sus ideas no han sobrevivido, pero otras siguen influyendo en el trabajo de los psicólogos modernos.

La psicología de la *Gestalt* se originó cuando, en 1912, Wertheimer, que estaba trabajando en el Instituto de Psicología de Frankfurt am Main, publicó un artículo sobre una ilusión visual denominada movimiento aparente. Trátase de la percepción de movimiento resultante de mirar una rápida secuencia de imágenes estacionarias, como ocurre en el cinematógrafo [*véase la figura 2]*. Este fenómeno indicó a Wertheimer que la percepción de un todo (el movimiento) era radicalmente diferente de la percepción de sus componentes (las imágenes estáticas).

La idea de que el todo es diferente de la suma de sus partes –tesis central de la psicología guestáltica– era un desafío a la entonces prevaleciente teoría del estructuralismo. En particular, los guestaltistas rechazaban el elementalismo, supuesto básico de los estructuralistas según el cual las percepciones complejas sólo podían entenderse identificando las partes elementales de la experiencia. Creían los estructuralistas que un observador bien entrenado podía llegar a distinguir los elementos básicos de la percepción y a reducirlos a sensaciones primitivas, tales como las de los puntos que constituyen un cuadrado o las de las diferentes notas de una melodía. Afirmaban que un cuadrado era precisamente la experiencia de un determinado haz de puntos que estimulaban la retina; y una melodía no era sino la experiencia de una serie de distintos tonos que se iban asociando entre sí en la mente de quien los oía. Su teorización se ha descrito como una especie de "química mental", pues da por supuesto que las percepciones pueden analizarse componente a componente, al igual que las moléculas pueden dividirse en átomos.

Los guestaltistas combatieron esta teoría. A su entender, lo que percibimos no es mera suma o serie de sensaciones, sino la configuración total de la que forman parte. La localización o el tamaño de la imagen de un cuadrado pueden alterarse de modo que se produzcan sensaciones retinianas enteramente diversas y, no obstante, la percepción sigue siendo todavía la de un cuadrado. ¿Y cómo se explica que pueda seguirse oyendo la misma melodía aunque se la trasponga a otra clave? Todos sus tonos quedan correlativamente modificados, y, sin embargo, sólo unos pocos músicos de finísimo oído advertirán el cambio.

Los teóricos de la *Gestalt* mantenían que las partes de un cuadrado –o las notas de una melodía– interactúan

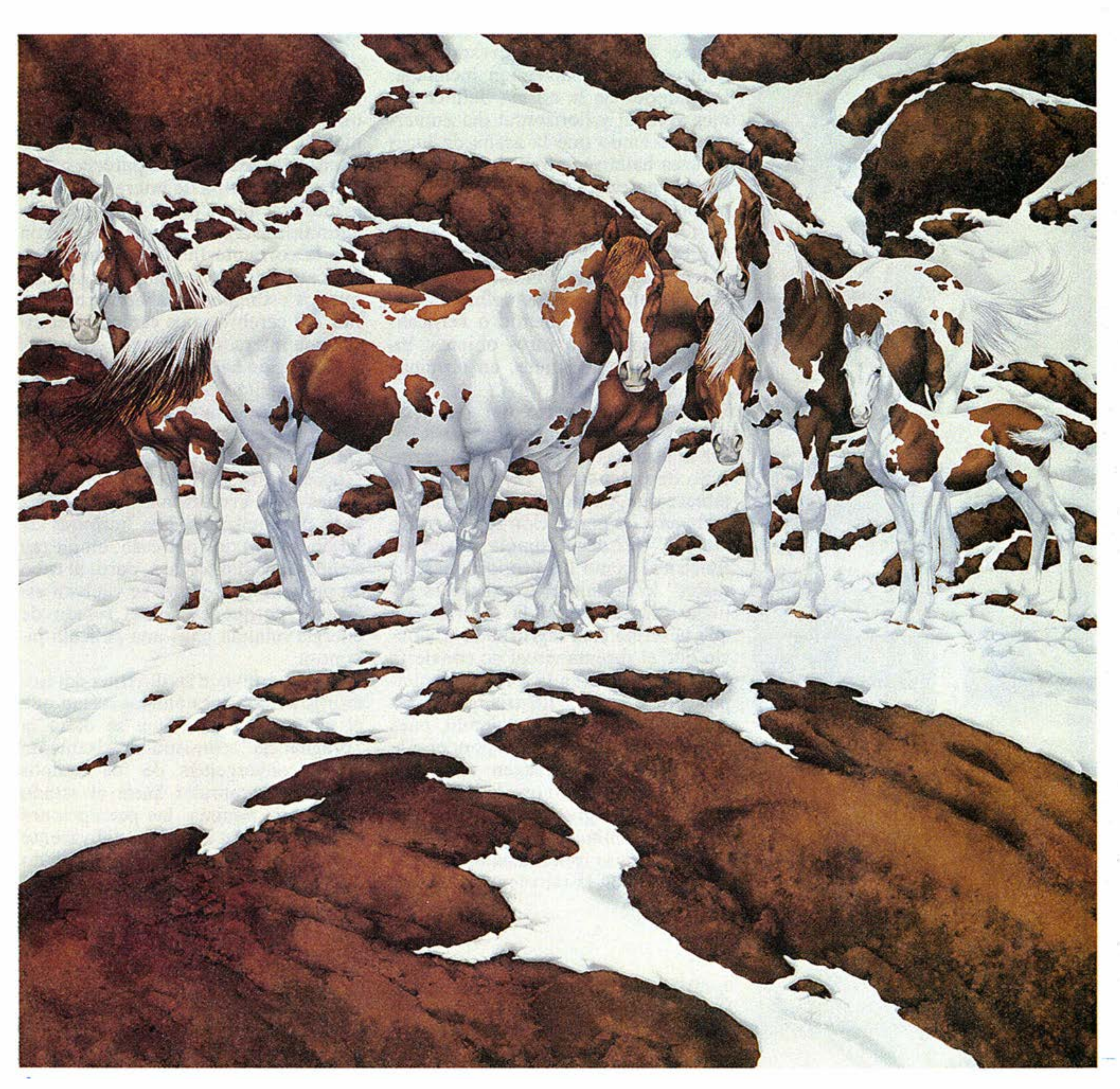
IRVING ROCK y STEPHEN PALMER, de la Universidad de California en Berkeley, investigan, en colaboración, la percepción visual. Aunque de distinta formación académica, comparten el interés por muchos fenómenos descubiertos por los psicólogos de la *Gestalt*. Rock fue alumno de la Nueva Escuela de Investigación Social, donde siguió las enseñanzas de discípulos directos de los fundadores de la *Gestaltpsychologie*, entre ellos Solomon Asch, Hans Wallach, Mary Henle y Martin Scheerer, hasta doctorarse allí, en 1952. Palmer se formó en la Universidad de California en San Diego, en la tradición más moderna del procesado informático, bajo la guía de Donald Norman y David Rumelhart. Su tesis doctoral, defendida en 1975, fue un intento de investigar las ideas de la *Gestalt* en términos informáticos. Rock y Palmer llevan adelante en la actualidad varios proyectos de investigación encaminados a ampliar y revisar las teorías guestálticas sobre la agrupación perceptual y el marco de referencia.

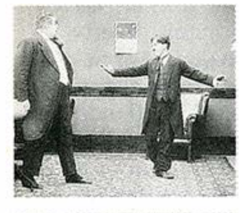
1. CAMUFLAJE NATURAL que muestra cómo determinadas leyes de agrupamiento (tales como la similaridad, la proximidad y la cohesión) pueden ocultar a unos animales. No resulta fácil distinguir esos poneys del paisaje circundante.

unas con otras y, con ello, hacen que se perciba un todo que resulta distinto de la mera suma de sus partes. La figura y la melodía constituyen ejemplos de lo que denominaron "propiedades emergentes": cualidades del conjunto de una experiencia que no son inherentes a cada uno de sus elementos. Ahora bien, las propiedades emergentes no se dan tan sólo cuando se trata de fenómenos mentales. Así, por ejemplo, las propiedades de la sal de mesa difieren mucho de las de sus componentes, el sodio (un metal corrosivo) y el cloro (un gas tóxico). También las características de una sociedad son distintas de las de los individuos que la componen.

La cualidad emergente ilustra un significado del concepto guestáltico de organización. Los guestaltistas creían además que la organización era necesaria para explicar por qué los seres humanos vemos el mundo como compuesto de objetos distintos. Ellos hicieron notar que, siendo la imagen retiniana tan sólo un conjunto de varias intensidades y frecuencias luminosas, los rayos procedentes de las distintas partes de un mismo objeto no tienen entre sí mayor afinidad que la que posean los que provienen de dos objetos diferentes. Por tanto, la capacidad de percibir objetos –piedras, árboles o casas– ha de deberse a una organización efectuada por el sistema nervioso. El caer en la cuenta de que la percepción de objetos separados no la producía por sí sola la "imagen" reflejada en la retina fue una de las más importantes contribuciones de los guestaltistas.

Para explicar cómo se forman las percepciones de los objetos, Wertheimer propuso la teoría de que el sistema visual organiza las partes según unas leyes de agrupación. Se tiende a agrupar perceptualmente aquellos elementos que están próximos unos a otros, se parecen entre sí, forman un contorno cerrado o se mueven en la misma dirección. La mayoría de las veces, estas leyes conducen a una representación exacta de los objetos que hay en la escena. También pueden dar lugar a representaciones confusas o menos nítidas, como en el caso del camuflaje. Otro aspecto importante de la percepción de imágenes, llamado organización figura-fondo, fue descubierto en 1921 por el psicólogo danés Edgar Rubin. Hizo éste notar que, aun cuando todas las par-







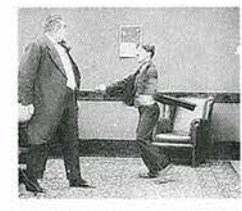


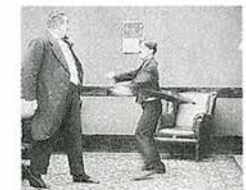










2. FOTOGRAMAS DE UNA PELICULA de Charlie Chaplin que patentizan cómo se crea la ilusión del movimiento aparente. Las imágenes estáticas dan la impresión de movimiento si se las mira, una tras otra, en rápida secuencia.

tes de una región conectada se agruparan adecuadamente, las tales podían interpretarse como objeto (figura) o como telón de encuadre (fondo). Estableció un conjunto de leyes que describen las condiciones en virtud de las cuales una región tiende a percibirse como figura y no como fondo.

Los guestaltistas descubrieron además que ciertas estructuras determinan un marco de referencia con respecto al cual se perciben otros objetos. Muchas personas han tenido alguna vez la experiencia de ese fenómeno llamado movimiento inducido, que se produce cuando el tren de al lado empieza a ponerse en marcha y confiere al observador la impresión de que el que se mueve, y en sentido opuesto, es el tren en que se halla sentado, por más que permanezca inmóvil. Otro ejemplo de tal fenómeno ocurre cuando un observador está dentro de una habitación inclinada: las paredes de la misma definen los ejes vertical y horizontal del enmarque, haciendo que la araña de luces parezca hallarse extrañamente ladeada, y lo parezca también el cuerpo del propio observador, pese a que una y otro estén perfectamente alineados con la gravedad. En ambos casos, el sistema visual usa una amplia estructura circundante para definir el patrón perceptual –quietud o verticalidad– e interpreta otros objetos, incluido el propio sujeto, en términos de esos patrones.

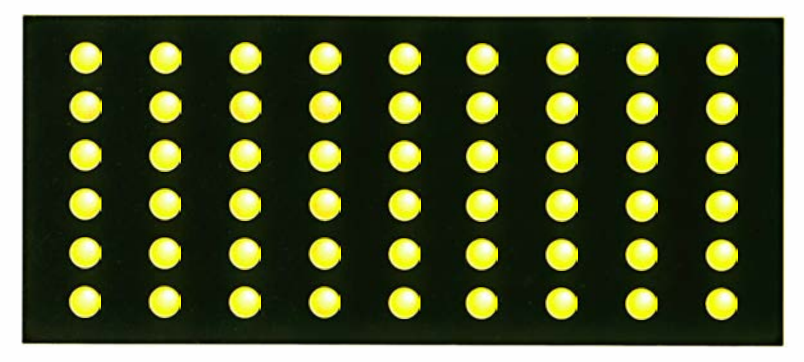
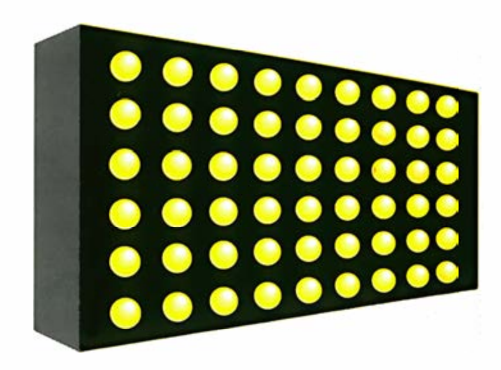
Un último aspecto del concepto guestáltico de organización versa sobre lo que sus autores llamaron "principio de pregnancia" (*Prägnanz*); establece que, cuando los estímulos son ambiguos, la percepción suele ser tan "buena" (es decir, simple, regular y simétrica) como las "condiciones actuales" lo permitan. Condiciones actuales significa aquí la información que la retina está registrando. Es obvio que el sistema visual no convierte cualquier figura en la forma más simple. Por ejemplo, un triángulo irregular no se ve como un círculo, pues la percepción ha de corresponder a la naturaleza de la imagen retiniana. Pero en los casos en que la imagen es ambigua, verbigracia, una figura parcialmente oculta, el observador tiende a percibir la forma más simple que concuerde con la información de que dispone.

Los teóricos de la *Gestalt* trataron de interpretar este y otros fenómenos perceptuales en términos fisiológicos. Postularon una conexión muy directa entre experiencia y fisiología en su doctrina del isomorfismo, que afirma que una experiencia subjetiva y el evento nervioso que a ella subyace tienen similares estructuras. El análisis del movimiento aparente verificado por Wertheimer ilustra esta idea. Cuando dos luces situadas una cerca de otra se encienden y apagan alternativamente a un ritmo apropiado, el observador ve una sola luz que se mueve de aquí para allá. Wertheimer sostenía que esta percepción debíase a la energía eléctrica que fluía entre las dos localizaciones cerebrales estimuladas por las luces; en otras palabras, que el episodio fisiológico tenía la misma estructura que la percepción por él producida.

Tal flujo de energía eléctrica por el cerebro no aludía a la transmisión de señales eléctricas de unas neuronas a otras, como lo impone la común opinión de los neurofisiólogos. Semejante sistema neuronal no parecía capaz de explicar el tipo de interacción y organización que concebían los teóricos guestaltistas, por lo que sugirieron que la corriente fluía directamente a través del tejido cerebral. Sostenían que los estímulos creaban en el interior del cerebro unos campos eléctricos que interactuaban unos con otros y convergían hacia un estado de energía mínima. Köhler, buen conocedor de la física de su tiempo, argüía que el cerebro era sólo un ejemplo de muchos sistemas físicos –llamados por él *Gestalten* físicas, "configuraciones" físicas– que evolucionan hacia un estado de equilibrio. Las burbujas de jabón, pongamos por caso, empiezan siendo de varias formas, pero, al cabo de un tiempo, siempre se vuelven esféricas, porque ese es el estado de energía mínima para una película jabonosa.

De acuerdo con su doctrina del isomorfismo, los guestaltistas creían que el mecanismo para que se diera la "pregnancia" consistía precisamente en la convergencia de los campos eléctricos cerebrales hacia el estado de energía mínima: las percepciones se simplificaban cuando el subyacente evento cerebral alcanzaba un estado de equilibrio.

Aunque nadie acepta ya la teoría de Köhler de los campos eléctricos del cerebro, otras muchas ideas que dimanaron de la psicología de la forma siguen influyendo actualmente en los teóricos de la percepción. En algunos casos, los puntos de vista de

3. RETICULO DE FOCOS LUMINOSOS organizado verticalmente en columnas (*panel izquierdo*). Cuando se le empuja hacia atrás por un lado (*panel derecho*), los observadores siguen percibiendo columnas, aunque las imágenes retinianas de los focos estén ahora más juntas horizontalmente.

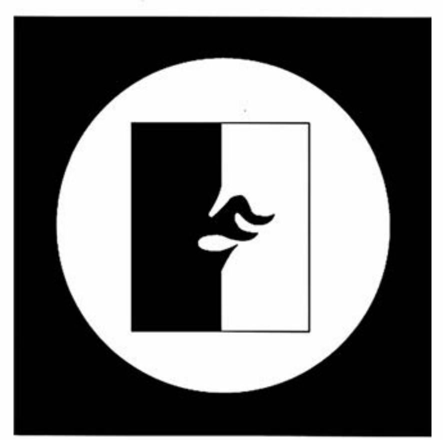
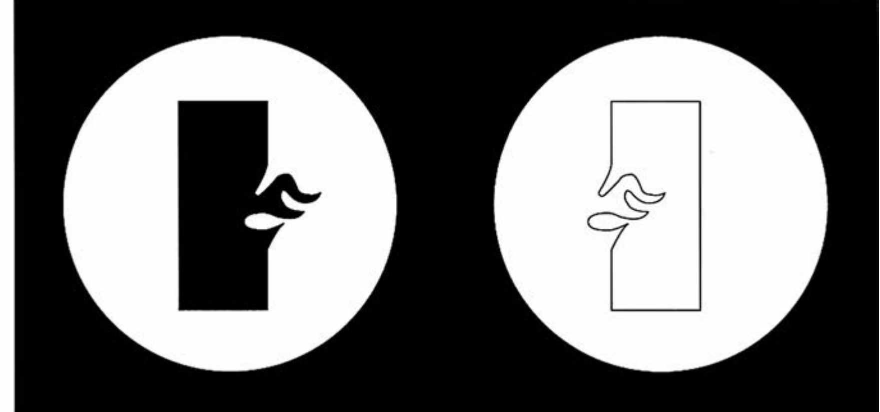
la *Gestalt* se han ampliado; en otros se han revisado, pero no hay un solo libro de texto actual sobre la percepción que no incluya gran número de ideas que tuvieron su origen en el movimiento guestáltico.

Las leyes de agrupación enunciadas por Wertheimer han resistido el paso del tiempo. En realidad, ninguna de ellas se ha visto refutada, ni tampoco se habían añadido otras nuevas a su lista original antes de las recientemente propuestas por nosotros. Uno de los autores (Palmer) sugirió una ley de inclusión, o de región común, refiriéndose a la tendencia de un observador a agrupar los elementos situados dentro de lo que se percibe como una misma región. La segunda ley, de cohesión, que postulamos conjuntamente, puede que sea el principio de agrupación más fundamental descubierto hasta la fecha. La "cohesión" se refiere a la fuerte tendencia del sistema visual a percibir toda región uniforme y conectada –por ejemplo un punto, una línea o un área más extensa– como una sola unidad. La cohesión constituye un candidato particularmente bueno para figurar como ley de agrupación, porque es quizá la propiedad más característica de los objetos que nos rodean. Sospechamos que a Wertheimer se le pasó por alto este importante principio al no caer en la cuenta de que había que explicar por qué cada elemento de sus configuraciones se percibía como una sola entidad.

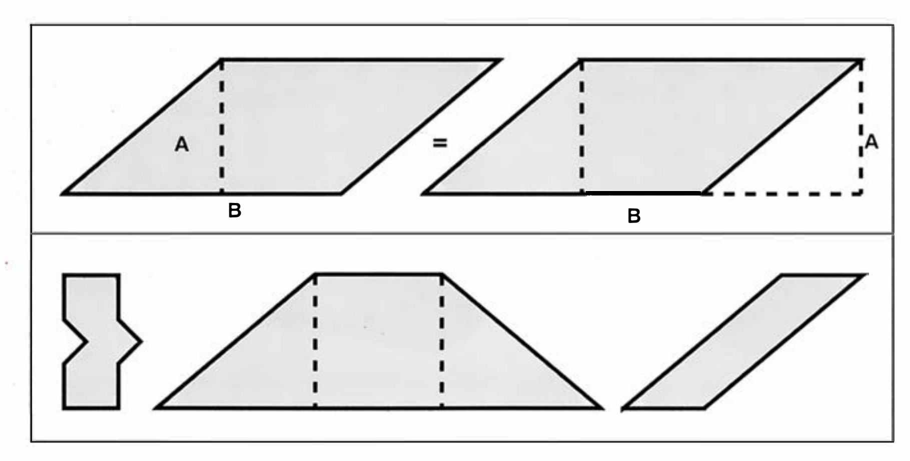
Aunque la validez de las leyes de agrupación no se ha puesto seriamente en duda, se está reconsiderando en qué fase del proceso visual operan. La tesis de la *Gestalt* da implícitamente por supuesto que la agrupación ha de ocurrir al principio del proceso visivo. Así, Wertheimer, al hablar del principio de proximidad, se refería presumiblemente a la proximidad retiniana: a lo inmediato que estuvieran unos de otros los estímulos en la retina. Sin embargo, es posible que los principios de agrupación operen en fases tardías del proceso, después de que se hayan percibido las condiciones del relieve y de la iluminación.

Para poner en claro estas dos hipótesis, uno de los autores (Rock) realizó hace algunos años un experimento con la colaboración de Leonard Brosgole. Ensartamos unos focos luminosos en unas cintas paralelas y suspendimos éstas en la oscuridad, de modo que el conjunto pareciese una retícula de puntos resplandecientes. Como los focos estaban más juntos unos a otros en el sentido vertical que en el horizontal, los observadores los veían organizados en columnas. Giramos luego el sistema hacia atrás, de modo que las imágenes retinianas de los focos se acercasen más unas a otras horizontalmente, aunque los focos mismos seguían estando, por supuesto, más juntos en sentido vertical. Cuando los observadores miraron esta imagen, siguieron viendo los focos en columnas, lo que indicaba que la agrupación se basaba en la proximidad percibida en el espacio tridimensional, más que en la que se diera de hecho en la retina. Por tanto, la agrupación por proximidad ha de ocurrir después de la percepción del relieve. A conclusiones parecidas hemos llegado respecto a los principios de región común y de cohesión, así como al del agrupar por similitud en la iluminación.

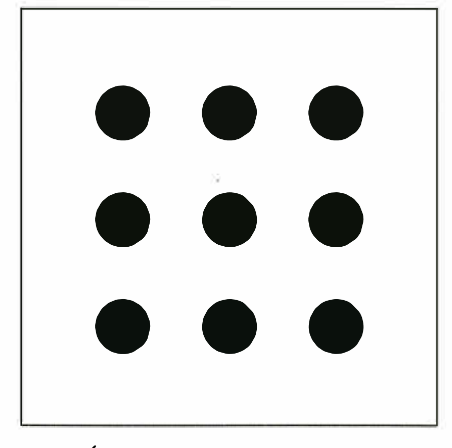
Nuevos métodos experimentales han permitido progresar en la comprensión del agrupamiento y han sugerido también que tiene vinculaciones con la fisiología subyacente. Jacob Beck, de la Universidad de

4. ORGANIZACION FIGURA-FONDO, fundamental para que se produzca la percepción. Cada lado de la imagen de la izquierda puede percibirse como figura o como fondo. Aunque las dos formas del panel de la derecha comparten el mismo perfil, parecen, sin embargo, muy diferentes.



5. EL COMPRENDER que el área de un paralelogramo es equivalente al área de un rectángulo (*en el recuadro superior*) hace más fácil hallar las superfícies de otras figuras (*en el recuadro inferior*). Un aprendizaje inteligente permite al observador transferir lo ya entendido a situaciones análogas.



6. CONÉCTENSE los discos negros trazando cuatro líneas rectas sin levantar el lápiz del papel. (La solución aparecerá en la figura 10).

Oregón, fue un adelantado en el estudio de la segregación textural, que es una manera de agrupar elementos por similaridad cuando se los percibe como una muestra, y no como formas individuales. En cierto experimento, ponía ante los observadores un campo con tres diferentes tipos de elementos yuxtapuestos: letras L (derecha o invertida), letras T y letras T inclinadas. Los observadores tenían que decir qué línea divisoria les parecía la más obvia o natural para distribuir en grupos la muestra.

Beck halló que la línea divisoria entre los elementos de tipo T derecha y los de tipo T inclinada era mucho más evidente que la que se podía trazar entre los de tipo L y los de tipo T. Lo cual revela –un tanto sorprendentemente desde el punto de vista guestáltico– que la orientación de los elementos es un factor más poderoso que la forma que tengan. Este y otros hallazgos afines han movido a los teóricos a suponer que hay conexiones entre la separación de las texturas y la actividad de las células del córtex visual que más responden a las diferencias en la orientación de las líneas componentes y de los bordes.

Otras técnicas han proporcionado modos de comprobar la tesis guestáltica de que los conjuntos son perceptualmente dominantes. David Navon, ahora en la Universidad israelita de Haifa, llevó a cabo un estudio para determinar si se percibían los todos antes que las partes o viceversa. Valiéndose de unas letras grandes compuestas de letras pequeñas, midió el tiempo que necesitaban los observadores para identificar las letras grandes ("globales") o las pequeñas ("locales"). En algunos casos, las letras grandes y las pequeñas eran las mismas (consistentes); en otros, diferían (estaban en conflicto).

Si el conjunto o todo es perceptualmente primario, según sostenían los guestaltistas, las letras globales deberán percibirse antes que las locales; si las partes priman en la percepción, como creen otros, deberá ser cierto lo contrario. Otra predicción guestáltica dice que, si primero se percibe el todo, las letras locales conflictivas no afectarán en nada al reconocimiento de las letras globales, mientras que unas letras conflictivas al nivel global retardarán el reconocimiento de las locales. Naturalmente, los teóricos que priman las partes prevén lo contrario. Pues bien, los resultados obtenidos por Navon en ambos casos corroboran las predicciones guestálticas. Con posterioridad, otros investigadores han visto que estos resultados no son tan convincentes como lo sugiere la teoría de la *Gestalt*, al demostrarse que las respuestas dependen del tamaño absoluto y relativo de las letras, entre otros factores.

Otro concepto de la teoría guestáltica que está aún muy vigente es el principio de "pregnancia" –la idea de que el sistema visual converge hacia la percepción más regular y simétrica coherente con la información sensorial. La vaga noción guestáltica de "bondad" se ha puesto en claro ahora. Cuando Emanuel Leeuwenberg y Hans Buffart trabajaban en la Universidad de Nimega, propusieron una teoría que especifica la cantidad de información en diversas percepciones –las "buenas" contienen poca información y las "malas" mucha–, y con ella predijeron cómo se percibirán figuras ocultas en parte, así como otros fenómenos. Wendell R. Garner, de la Universidad de Yale, ha demostrado que las formas buenas pueden captarse más deprisa, recordarse mejor y describirse con más concisión que las malas.

En contraste con sus teorías acerca de la percepción, las ideas de los guestaltistas sobre campos eléctricos en el cerebro han sido resueltamente rechazadas por los fisiólogos modernos. Sin embargo, conceptos similares a las nociones köhlerianas de *Gestalten* físicas han vuelto a aparecer bajo la forma de redes neuronales. Según los teorizadores de esas redes, los procesos mentales constituyen el resultado del comportamiento dinámico de muchas unidades interconectadas que pueden concebirse como neuronas. El comportamiento de cada una de esas unidades puede caracterizarse por su estado de activación –algo así como la velocidad de disparo de una neurona– y las unidades influyen unas en otras mediante conexiones excitadoras o inhibidoras –algo así como las sinapsis. El sistema entero se activa inicialmente con un estímulo externo que afecta a algún subconjunto de las unidades. La activación se propaga luego por toda la red hasta que se logra un estado equilibrado de energía mínima. En suma, esas redes nerviosas pueden concebirse como ejemplos de *Gestalten* físicas. Aunque en estos trabajos se están dando todavía los primeros pasos, los modelos neurorreticulares de la percepción prometen abrir un capítulo enteramente nuevo de la teoría de la *Gestalt*.

Además de revolucionar los estudios de la percepción, los teóricos guestaltistas hicieron valiosas aportaciones a los campos del aprendizaje, de la memoria y del pensamiento –con importantes implicaciones para la educación–, así como a la psicología social. Las primeras ideas guestálticas acerca del pensamiento chocaron con las del ascendente movimiento behaviorista. Un precursor

de esta escuela, Edward L. Thorndike, sacó de sus estudios la conclusión de que los animales resolvían problemas a través del método del ensayo y error, y no porque pensaran o entendieran. En un experimento hoy famoso, introducía a un gato en una jaula de la que el animal podía escapar si estiraba una cuerda que descorría el pestillo de la puerta. En el curso de sus idas y venidas por el interior de la jaula, el gato solía tirar alguna vez inadvertidamente de la cuerda y quedaba libre. Tras muchas de estas pruebas, pasaba a tirar de la cuerda al momento mismo de habérsele vuelto a encerrar en la jaula. Thorndike concluyó que el gato no usaba la inteligencia, sino que había ido desarrollando gradualmente una asociación.

Los teóricos de la *Gestalt* clamaron con vehemencia contra este tipo de experimento y contra las conclusiones deducidas del mismo. Objetaban que aquella situación impedía, de hecho, toda muestra de inteligencia en la resolución del problema, pues no era muy de esperar que el gato entendiese el oculto mecanismo que relacionaba el tirar de la cuerda con el abrirse la puerta. En cambio, Köhler llevó a cabo experimentos con chimpancés mientras estuvo aislado en la isla de Tenerife, durante la Primera Guerra Mundial, y en ellos se hicieron perceptualmente evidentes tanto los requisitos para llegar a una solución como los medios de lograrla. Köhler observó que los monos descubrían la manera de alcanzar y coger, utilizando un palo, las bananas que veían fuera de la jaula en que estaban encerrados.

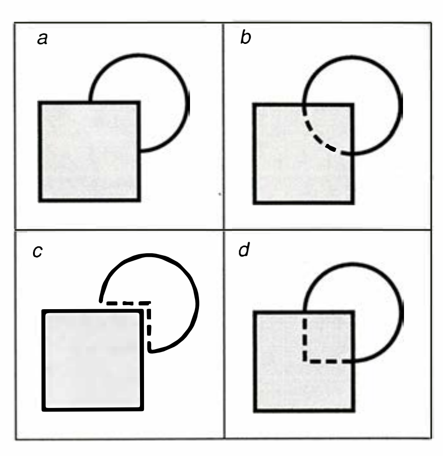
Estas comprobaciones entraban en conflicto con los dogmas behavioristas al menos en dos puntos importantes. Primero: los chimpancés daban con la solución de repente, en un instante de "comprensión", y no gradualmente. Esto era posible –argüía Köhler– porque la naturaleza del problema era perceptualmente manifiesta, a diferencia de lo que ocurría con el mecanismo cuerda-cerrojo de la puerta. Segundo: los errores cometidos por los monos no fueron fortuitos, como lo predecía la teoría behaviorista, sino que suponían un despliegue de inteligencia y comprensión.

Aunque nadie ha podido explicar cómo se produce esa especie de instantánea intelección, los guestaltistas aclararon sin duda ciertos aspectos de cómo podría lograrse. Un modo de conseguirla los humanos, a diferencia de los animales irracionales, es contar con alguien que les explique lo que han de entender. El mero escucharle no basta, desde luego, pues el oyente se ha de hacer con la misma estructura cognitiva del que ofrece la explicación, para enterarse de las conexiones esenciales entre los datos pertinentes. Eso no significa que los oyentes tengan que pasar por el mismo proceso creativo por el que pasó el solucionador original del problema para llegar a resolverlo; pero su estado final de comprensión sí que ha de ser parecido.

No se pueden sobrestimar las implicaciones educativas del conseguir comprender mediante la explicación. No sólo es gratificante dar así con la solución de un problema, sino que es mucho menos probable su olvido que si se la aprende por rutinaria memorización; y, además, se la podrá transferir fácilmente a nuevos problemas semejantes. Wertheimer demostró, por ejemplo, que una vez caen los niños en la cuenta de por qué el área de un paralelogramo equivale a su base multiplicada por su altura, pueden hallar las áreas de otras figuras geométricas sin tener que memorizar las fórmulas. Muchos educadores modernos que critican el aprendizaje rutinario abogan por que se enseñe a los estudiantes a pensar creativamente y a lograr entender. Pocos son los que advierten que estas "revolucionarias" ideas acerca de la educación tuvieron su origen en los psicólogos de la *Gestalt*.

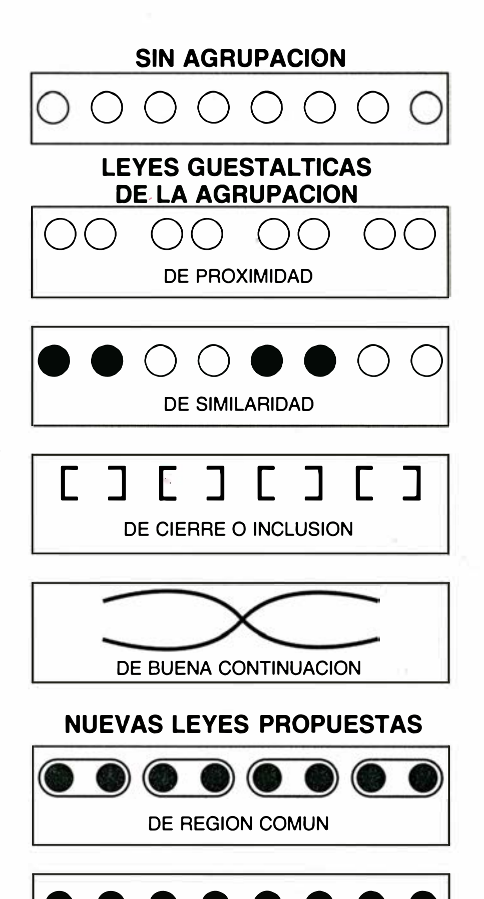
Los teóricos guestaltistas se esforzaron también por describir el proceso creativo mediante el cual consigue una persona la genuina comprensión en la vida de cada día. Según ellos, los problemas tienen ciertas exigencias fáciles de captar, que llevan a la gente a procurar solucionarlos de manera no fortuita. El aferrarse a una hipótesis o a una función de un objeto –a menudo sin darse cuenta de ello– es el principal obstáculo para la intelección. Cuando uno se desembaraza de suposiciones implícitas, su manera de ver un problema se reorganiza a veces por completo, capacitándole de pronto para "ver" la solución, lo que suele ir acompañado de un "¡ajá!" de satisfacción.

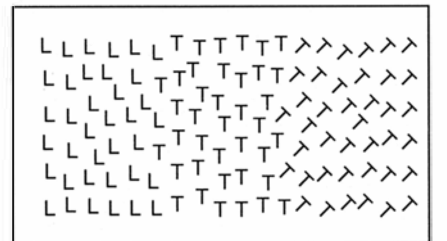
Los consagrados a la investigación de cómo solucionamos problemas los humanos no han podido explicar aún el hecho del entender, pero han abandonado la idea behaviorista de que consiste en ciegos tanteos y errores y se pronuncian a favor de otra más en consonancia con las concepciones guestálticas sobre el valor de la comprensión. Un punto sobre el que se concentra prometedoramente la investigación reciente es el del uso de analogías en la resolución de problemas: quienes entienden un asunto pueden aplicar ese conocimiento a cualquier otro mediante analogías.



7. LA FIGURA OSCURA ilustra la idea de "pregnancia". Dada una configuración ambigua (*a*), los observadores perciben formas simples (*b*), en vez de formas complejas (*c*, *d*).

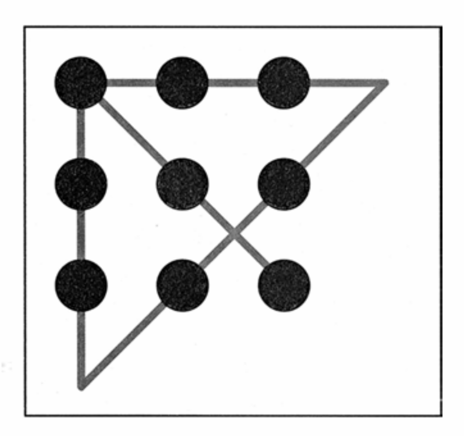
Los psicólogos guestaltistas hicieron ulteriores incursiones contra el enfoque behaviorista en el terreno de la psicología social. Para empezar, a finales de la década los treinta, tres investigadores –Kurt Lewin, Fritz





8. SEGREGACION TEXTURAL: fundada en la disimilaridad de los elementos. La línea de separación entre las letras T derechas y las ladeadas es más clara que la divisoria que separa las letras T derechas y las letras L *(izquierda).*

9. EN VIRTUD DE LA PRECEDENCIA de la forma global: podemos reconocer las letras mayores más rápidamente que las pequeñas, sean éstas coherentes o no. En cambio, la identificación de las letras pequeñas lleva más tiempo cuando están en conflicto con la letra global.



10. LOS DISCOS PUEDEN CONECTARSE entre sí trazando las líneas más allá de los propios círculos. Se equivocan, pues, quienes dan por supuesto que no se puede hacer la unión.

Heider y Solomon E. Asch– rechazaron la idea de que el comportamiento social pudiera explicarse tan sólo como una respuesta condicionada por recompensas sociales, como la aprobación o las alabanzas. Antes bien, argüían, la gente comprende la conducta de los demás atribuyéndoles sentimientos, sensibilidad, aspiraciones, creencias e intenciones –opinión ésta conocida como teoría de la atribución. Por muy obvia que parezca esta idea, suponía un abandono radical del enfoque behaviorista, en boga a la sazón, que minimizaba o negaba los estados mentales subjetivos. La teoría de la atribución ha desplazado desde entonces el behaviorismo como opinión predominante en psicología social.

Pocas de las ideas de Lewin han sobrevivido en la psicología contemporánea, mientras que la obra de Heider y Asch ha ejercido una influencia duradera. Heider aplicó las concepciones guestálticas sobre la percepción de un objeto a la percepción de las demás personas. Una piedra angular de su teoría era la idea de atribución: que la gente trata de explicarse la conducta de unos y otros por relaciones causales más profundas, tales como motivos e intenciones, apelando a coherencias contextuales y de comportamiento. Heider desarrolló también la idea del equilibrio compensatorio, la concepción de que los individuos prefieren las relaciones cognitivas armoniosas. Así, por ejemplo, si a Juana le gusta *X* y piensa que a éste le gusta a su vez *Y*, el sistema de opiniones quedará equilibrado si a Juana le gusta también *Y*, y desequilibrado o descompensado si a ella no le gusta *Y*. Esta idea es como un eco del principio de pregnancia: hay tendencia a lograr la organización mejor o más fundamentada.

El trabajo de siembra que efectuó Heider con su teoría del equilibrio está relacionado con la teoría del desacuerdo cognitivo, que ideó posteriormente Leon Festinger. Persuadido éste de que la gente procura reducir las incoherencias en sus opiniones, sentimientos y actitudes, se dedicó a estudiar el influjo de las opciones de la gente en sus subsiguientes opiniones y actitudes. Razonó que, cuando la alternativa rechazada (digamos, un coche deportivo pero de funcionamiento poco seguro) es, en muchos aspectos, más deseable que la elegida (un coche menos llamativo aunque más de fiar), el hecho de no ser escogida suele traer consigo un estado interior de desarmonía –o desacuerdo, como lo llamó Festinger– que produce una presión orientada a eliminarlo. Un modo de reducir el desacuerdo es revalorar los atractivos de las alternativas, por ejemplo devaluando la desechada (a fin de cuentas, los coches deportivos son demasiado peligrosos) y dando, con ello, superior valor a la escogida.

Asch, que trabajó con Wertheimer en la Nueva Escuela de Investigación Social, extendió directamente la teoría guestáltica a la psicología social. Sostuvo que las actitudes están enraizadas en las creencias y, éstas, en la información; manifestó también que las opiniones tienden a ser racionales más que a dejarse moldear por la "sugestión", como lo pensaban los primeros psicólogos sociales. Su insistencia en la racionalidad humana contrastaba mucho con la aparente irracionalidad del prejuicio racial y otros fenómenos similares. Asch argüía, sin embargo, que hasta ese prejuicio se puede entender como algo razonable y enraizado en la información, aunque en el caso sea una información errónea. Por ejemplo, si los niños dependen de sus padres y de otros adultos respetados y tienen pocos motivos para desconfiar de éstos, es razonable que acepten como buenas las opiniones de esas personas mayores respecto a un grupo étnico o una raza. Además, los niños pueden obtener poca, si es que alguna, información de otras fuentes para contradecir lo que sus padres les digan.

Asch cuestionó también el supuesto behaviorista según el cual las creencias y las actitudes son el resultado de sugestiones basadas en el prestigio de la fuente informante. Por ejemplo, era sabido que los estudiantes de enseñanza secundaria estadounidenses modificaban su opinión acerca de una frase según quién creyeran que la había pronunciado. Cuando se les aseguraba que Thomas Jefferson había dicho que "una pequeña rebelión es, de vez en cuando, buena cosa, y tan necesaria en el mundo político como las tormentas en el físico", asentían a menudo con firmeza. Si la misma frase se atribuía a Lenin, su acuerdo con tal afirmación disminuía considerablemente.

Estos datos respaldan a primera vista la idea de que la aceptabilidad de la fuente influye mucho en las opiniones de una persona. Pero Asch creía que tales decisiones tienen un fundamento racional. Sostenía que la gente interpretaba diferentemente la afirmación según quién pensara que la había hecho. En el caso en cuestión, Asch averiguó que los estudiantes que atribuían esa frase a Lenin interpretaban "rebelión" como total trastorno del sistema socioeconómico; en cambio, los que la atribuían a Jefferson solían dar al término una interpretación menos violenta, entendiendo por él una moderada reforma social o política.

Este aspecto de la obra de Asch representa una ampliación del concepto guestáltico de los efectos contextuales parte-todo, que fue desarrollado para

explicar fenómenos perceptuales. En este caso, la parte era la afirmación, la cual venía a tener significados diferentes cuando se insertaba en el conjunto (todo lo que uno sabe acerca del autor de la frase, sea éste Jefferson o Lenin). La idea del todo resultante de la organización de partes se aclara en el experimento que hizo Asch acerca de cómo la gente se forma impresiones de las personalidades a partir de unas listas de rasgos. Comprobó Asch que, cuando alguien llega a tener una impresión global de una persona, ciertos rasgos son esenciales: sustituyendo unos por otros en una lista de rasgos, por lo demás idénticos, suele cambiar enteramente la impresión general. Aun el mismo rasgo se percibirá de manera diferente en función de otro rasgo. Así, el significado de ser una persona "decidida" no es del todo igual tratándose de un sujeto ardoroso que tratándose de un sujeto frío.

En algunos aspectos, el movimiento de la *Gestalt*, pese a su reconocido influjo en varios campos de la psicología, siempre se ha mirado con bastante escepticismo desde la ciencia oficial. Este escepticismo era lógico allá por los años veinte y treinta, cuando estaban en boga las teorías del estructuralismo y del behaviorismo, escuelas contra las que los guestaltistas combatieron con vehemencia y con éxito. Sin embargo, tal escepticismo persiste todavía hoy, por varias razones. En primer lugar, la psicología de la forma trató de investigar la experiencia subjetiva, por ejemplo en la percepción, lo cual era rechazado por los behavioristas como objeto impropio para la investigación científica. En segundo lugar, aunque los guestaltistas llevaron a cabo muchos experimentos bien controlados, sus fenómenos mejor conocidos los presentaban a menudo como demostraciones directas, haciéndolo, por ejemplo, con las figuras con que ilustraban las leyes de agrupación. En tercer lugar, sus teorías solían expresarse en términos cualitativos, sin ajustarse a los patrones de precisión corrientemente admitidos. Además, sus opiniones acerca del funcionamiento del cerebro se han visto desacreditadas por los modernos neurofisiólogos. Y por último, aunque no lo menos importante, el enfoque teórico que propugnaban parece negar uno de los principios básicos del método científico: el de que los todos pueden entenderse mediante su reducción a un conjunto de partes.

Estos formidables obstáculos que se oponen a la aceptación de las ideas de la *Gestalt* deberán contrapesarse con sus considerables logros. Es impresionante la lista de importantes fenómenos perceptuales que los psicólogos de esta escuela contribuyeron a dilucidar: agrupación, organización figura-fondo, marcos de referencia, bondad configuracional y movimiento aparente, para no citar más que los mentados en este artículo. Aunque sería lógicamente posible que tales descubrimientos se hubiesen hecho con independencia de los métodos y teorías guestálticos, parece improbable. El ataque de la *Gestalt* contra el estructuralismo tuvo efectos devastadores.

Por añadidura, los guestaltistas vencieron a los behavioristas en su controversia respecto a la naturaleza del aprendizaje, el pensamiento y la psicología social. Aunque actualmente los psicólogos siguen métodos behavioristas, la teoría del behaviorismo ha sido abandonada en favor de un enfoque cognitivo más en línea con la concepción guestáltica. Los problemas teóricos que, en torno a la organización perceptual, a la intelección, al aprendizaje y a la racionalidad humana, plantearon los psicólogos de la *Gestalt* siguen siendo de los más profundos y más complejos que se dan en psicología.

Aunque las ideas guestálticas acerca de los campos eléctricos del cerebro eran erróneas, la tesis más general de que el cerebro es un sistema dinámico que converge hacia el equilibrio en su funcionamiento energético –*Gestalten* físicas en la terminología de Köhler– puede ser correcta. El notable auge del interés por los modelos de redes y circuitos neuronales atestigua que las teorías de la *Gestalt* siguen teniendo hoy mucha validez y ocupan un lugar seguro en la historia de la psicología.

BIBLIOGRAFIA COMPLEMENTARIA

A SOURCE BOOK OF GESTALT PSYCHOLOGY. Preparado por Willis Davis Ellis. Humanities Press, 1938.

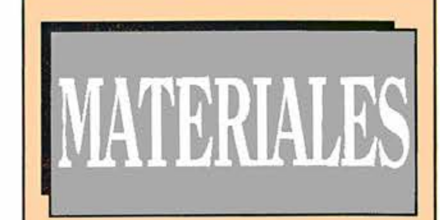
GESTALT PSYCHOLOGY. Wolfgang Köhler. Liveright, 1970.

THE MENTALITY OF APES. Wolgang Köhler. Liveright, 1976.

PERCEPTUAL ORGANIZATION. Dirigido por Michael Kubovy y James R. Pomerantz. Lawrence Erlbaum Associates, 1981.

PARALLEL DISTRIBUTED PROCESSING: EXPLORATIONS IN THE MICROSTRUCTURE OF COGNITION, Vol 1: FOUNDATIONS. David E. Rumelhart, James L. McClelland y PDP Research Group. The MIT Press, 1986.

SOCIAL PSYCHOLOGY. Solomon E. Asch. Oxford University Press, 1987.



Congelados y vivos

Diversos animales se congelan durante los meses de invierno y se deshielan en primavera. Esta capacidad natural para sobrevivir a la congelación puede servirnos de guía para abordar la criopreservación de tejidos humanos

Kenneth B. Storey y Janet M. Storey

Cuando el mercurio desciende por debajo de los cero grados Celsius, nos refugiamos en nuestros cálidos hogares, nos protegemos con afelpados anoraks si hemos de salir a la calle y pensamos en lo bien que nos vendrían unas vacaciones en algún lugar tropical. No abundan los animales que siguen activos durante los meses de invierno. Las aves han volado hacia el sur; muchos animales de hábitos terrestres hibernan en madrigueras o en el fondo de los lagos, si acuáticos. Pero, ¿qué les ocurre a los animales ectotérmicos, o de sangre fría (ranas y tortugas, escarabajos y arañas), que no pueden encontrar refugio medianamente cálido? ¿Cómo resisten cuando la temperatura cae por debajo del punto de congelación de sus fluidos corporales? Algunas especies evitan la congelación mediante cambios bioquímicos en su cuerpo. Curiosamente, otras muchas reaccionan congelándose para sobrevivir.

Cientos de especies de insectos terrestres sobreviven tras largos períodos de congelación invernal. Por citar casos extremos, ciertos insectos que viven en pleno Artico, como las orugas del oso lanudo (*Gynaephora groenlandica*), pasan 10 meses del año congeladas a temperaturas que descienden a 50 grados centígrados bajo cero o incluso inferiores. Varios animales invertebrados que colonizan la zona intermareal de las costas marinas septentrionales, como bellotas de mar, mejillones y bígaros, se congelan también cuando se hallan expuestos a temperaturas del aire bajo cero en marea baja. Para los propósitos de nuestro laboratorio de la Universidad de Carleton en Ottawa, reviste mayor interés un grupo de anfibios y reptiles que sobreviven a la congelación durante su hibernación.

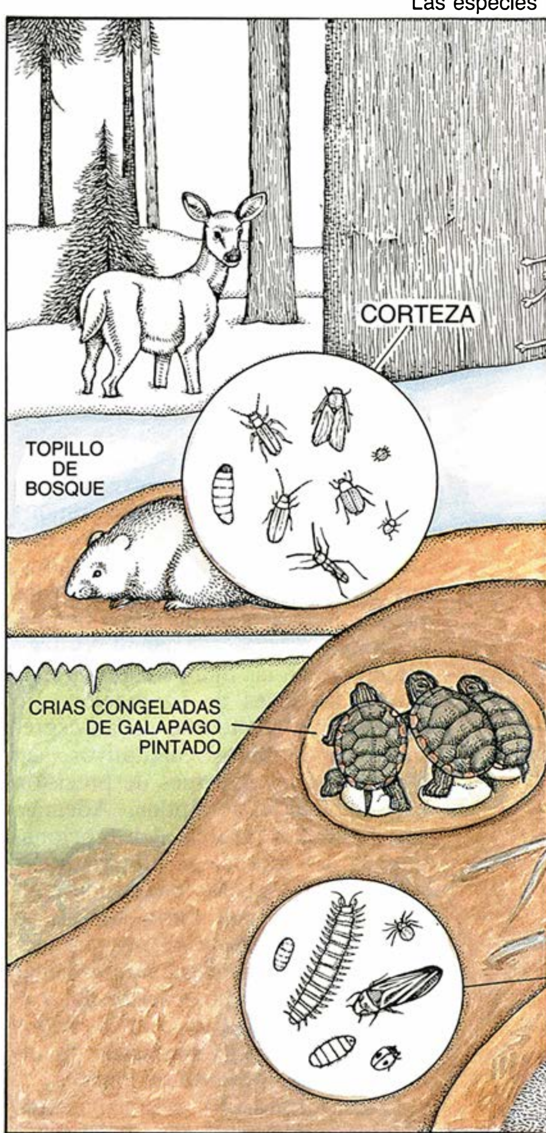
A William D. Schmid, de la Universidad de Minnesota en Minneapolis, debemos nuestras investigaciones sobre esa sorprendente adaptación. En 1982 publicó un informe sobre las ranas que sobreviven a la congelación, que dio pie a nuestro trabajo. Continuando los estudios de Schmid, hemos demostrado que cuatro especies comunes de ranas (la rana de bosque, *Rana sylvatica*; la ranita arbórea, *Hyla crucifer*; la ranita arbórea gris, *Hyla versicolor*, y la ranita listada, *Pseudacris triseriata*), que hibernan en el suelo del bosque, pueden sobrevivir días o semanas de congelación con hasta el 65 por ciento del agua total de su cuerpo convertida en hielo. Científicos soviéticos han señalado a su vez que la salamandra siberiana (*Hynobius keyserlingi*) sobrevive a la congelación. Esta especie, el único anfibio que se encuentra en la tundra que hiberna en tierra, puede sobrevivir a la exposición a 35 grados C bajo cero.

En 1988, identificamos a un reptil que se congela durante el invierno. Nuestro colega Ronald J. Brooks, de la Universidad de Guelph, nos refirió el insólito comportamiento de los galápagos pintados (*Chrysemys picta*) acabados de salir del huevo. En vez de abandonar sus nidos después de hacer eclosión a finales de verano, las tortuguitas permanecen allí, a buen recaudo de los depredadores, hasta la llegada de la primavera. Sus nidos, situados entre sólo 7 y 10 centímetros de profundidad en las orillas expuestas de lagos y ríos, ofrecen escaso aislamiento.

Ronald Brooks registró temperaturas en el nido de 6 a 8 grados cen-

KENNETH B. STOREY y JANET M. STOREY, marido y mujer, exploran las adaptaciones bioquímicas de los animales en la Universidad de Carleton, en Ottawa. Kenneth Storey se licenció por la Universidad de Calgary y se doctoró por la de Columbia Británica. Miembro de la Sociedad Real de Canadá, es profesor de bioquímica en Carleton. Janet Storey se licenció en la Universidad de Manitoba y se graduó en la de Columbia Británica. El matrimonio comparte horas de investigación y de educación de sus dos hijas, quienes se han acostumbrado a ver frascos con ranas en el frigorífico de la cocina.

ESTRATEGIAS
Las especies



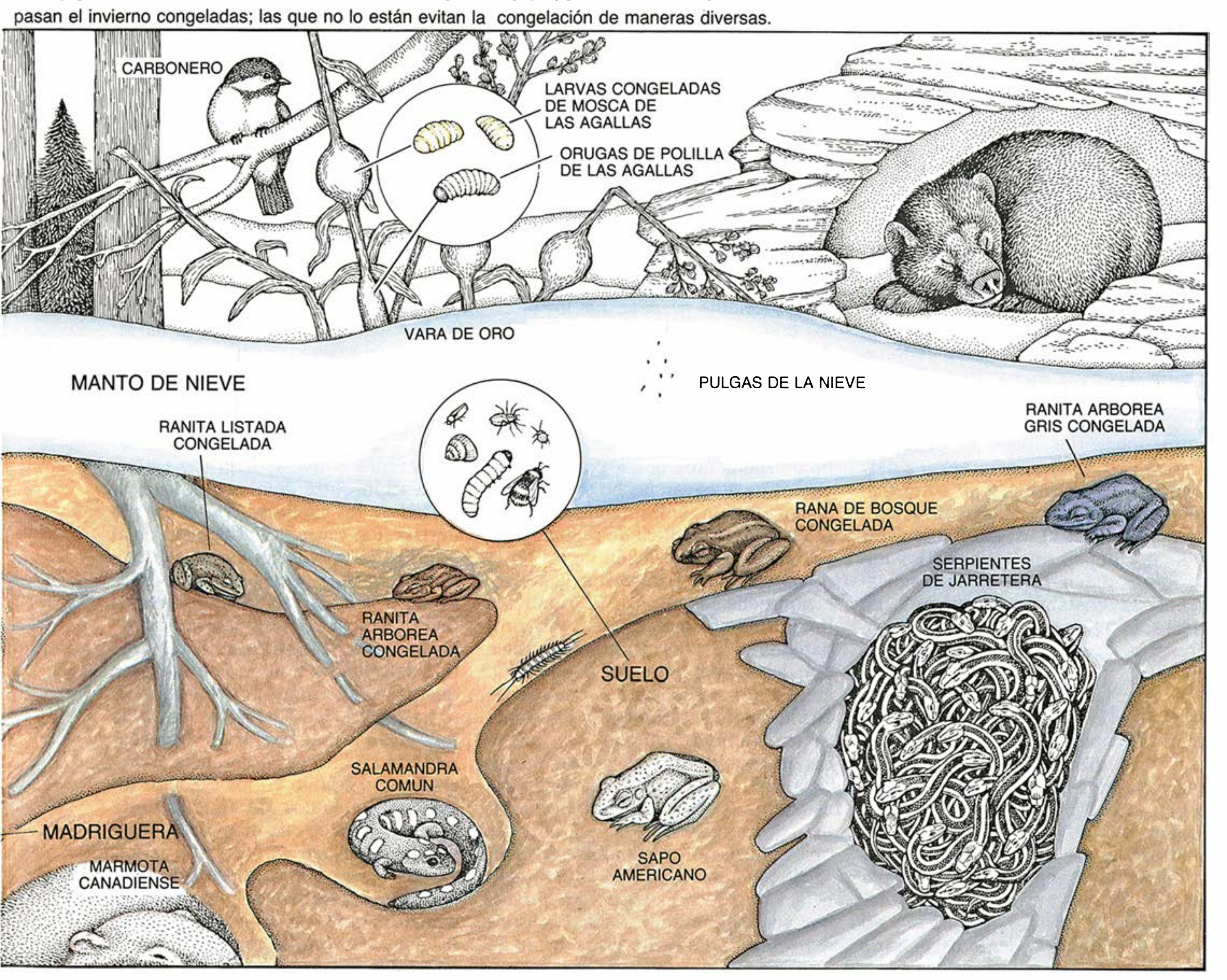
tígrados bajo cero durante enero y febrero de 1988, pero las pruebas de nuestro laboratorio mostraron que los galápagos se congelan siempre que la temperatura desciende por debajo de los —3 grados C. Por tanto, las tortuguitas han de congelarse y descongelarse varias veces a lo largo del invierno antes de emerger en primavera. Estudios realizados por Jon P. Costanzo, Dennis L. Claussen y Richard E. Lee, Jr., de la Universidad de Miami en Oxford, Ohio, mostraron también que las tortugas caja adultas y las serpientes de jarretera tienen cierta capacidad de sobrevivir a la congelación.

Mientras viven congelados, ninguno de estos animales se mueve ni presenta señales de respiración, latidos cardíacos o circulación sanguínea. A tenor de nuestros últimos experimentos, la actividad neurológica que desarrollan apenas es detectable. El hielo se acumula en todos los compartimentos fluidos extracelulares y ocupa la cavidad abdominal y la vejiga; bajo la piel y entre los músculos se encuentran cristales. Estos animales son maestros en los trucos de la criopreservación de órganos (la congelación de tejido vivo para su almacenamiento y uso posterior). Nuestros estudios de ranas y galápagos congelados están revelando los mecanismos moleculares esenciales para la vida en estado congelado.

Pasar el invierno en estado de congelación parece ser una estrategia adaptativa que encierra un notable peligro: la congelación es letal para la mayoría de células. Como sabe cualquier jardinero, la primera helada fuerte transforma un lozano campo florido otoñal en una masa pardusca. Los cristales de hielo crecen a través de las membranas celulares y dañan los orgánulos subcelulares; se derrama el contenido celular, y la localización discreta de los procesos metabólicos individuales en el interior de la célula se desordena. Aun cuando la formación de hielo pueda controlarse, la congelación sobrecarga a las células de otras maneras. Por ejemplo, como la congelación detiene la respiración y la circulación de la sangre, todos los órganos ven impedido el acceso al oxígeno y a los combustibles que aporta la sangre mientras dura la congelación. En vez de tolerar el estado congelado, ¿no hay maneras más fáciles mediante las cuales puedan los animales de sangre fría soportar temperaturas inferiores a cero?

La verdad es que hay dos estrategias alternativas. La primera, la más familiar también, consiste en evitar la exposición a temperaturas por debajo del punto de congelación de los fluidos corporales. Los animales simple-

DE SUPERVIVENCIA INVERNAL EN UN BOSQUE DE LA ZONA TEMPLADA
pasan el invierno congeladas; las que no lo están evitan la congelación de maneras diversas.



mente “eligen” lugares de hibernación bastante cálidos, bajo el agua o a cierta profundidad bajo tierra. Numerosas especies de insectos pasan el invierno en fase larvaria acuática; muchos tipos de ranas y galápagos hibernan en el fondo de estanques, donde se encuentran a salvo a menos que la masa de agua se hiele por completo. En tierra, los sapos pueden excavar en el suelo para permanecer por debajo del horizonte de helada, y las serpientes pueden congregarse en cubiles subterráneos comunales.

La segunda alternativa a la congelación consiste en utilizar adaptaciones específicas que estabilicen el estado líquido a temperaturas inferiores a cero. Todas las soluciones acuosas, entre ellas los fluidos corporales, tienen un punto de congelación de equilibrio, que corresponde a la temperatura a la que un cristal de hielo colocado en la solución empieza a crecer. Pero todas las soluciones acuosas pueden ser también sobreenfriadas; es decir, pueden enfriarse muy por debajo del punto de congelación de equilibrio sin que por ello el agua cristalice espontáneamente en forma de hielo. Así, el plasma humano tiene el punto de congelación a −0,8 grados C; ahora bien, cuando lo enfriamos de forma controlada, puede ser sobreenfriado hasta −16 grados C.

Sin embargo, la presencia de nucleadores limita la extensión del sobreenfriamiento. Llámanse nucleadores los compuestos que provocan el crecimiento del hielo al proporcionar puntos de enlace que ordenan las moléculas de agua en la estructura reticular del hielo. El propio hielo es el mejor nucleador, pero las proteínas plasmáticas, bacterias extrañas y partículas alimenticias actúan asimismo como eficaces nucleadores. Por ello, para estabilizar el estado líquido, los animales tienen que eliminar los nucleadores o evitar que éstos desencadenen la cristalización general; en efecto, han de reducir el punto de sobreenfriamiento de sus fluidos corporales.

Arthur L. DeVries y sus colegas, de la Universidad de Illinois en Urbana-Champaign, descubrieron que los peces marinos polares emplean una estrategia de ese tipo. Los peces evitan la congelación porque poseen proteínas anticongelantes en sus fluidos corporales. Cuando en el interior del pez comienzan a formarse cristales de hielo, las proteínas se unen rápidamente al cristal para impedir así la adición ulterior de moléculas de agua al plano de crecimiento del cristal [véase “Peces antárticos”, por Joseph T. Eastman y Arthur L. DeVries; Investigación y Ciencia, enero de 1987]. Muchos artrópodos terrestres, entre los que se cuentan arañas, garrapatas, ácaros y numerosos insectos, han desarrollado sus propias proteínas anticongelantes. En muchos casos, las proteínas anticongelantes de insectos son tan potentes que pueden impedir la formación de hielo a temperaturas muy bajas, de hasta −15 grados C, lo que permite que muchos insectos permanezcan activos bajo el manto de nieve invernal.

Otros insectos requieren una mayor protección invernal y han desarrollado un anticongelante adicional, además de las proteínas, para reducir el punto de sobreenfriamiento. Estos insectos cargan sus fluidos corporales con un anticongelante de azúcares, de polihidroxialcoholes de bajo peso molecular. Una solución de etilenglicol al 50 por ciento, que proporciona protección hasta los −30 grados C, es el nivel normal de anticongelante que se añade a los radiadores de los automóviles en el Canadá meridional. En comparación, nuestros estudios de las orugas de la polilla de las agallas (*Epiblema scudderiana*) demostraron que tienen fluidos corporales que en pleno invierno están constituidos por cerca de un 40 por ciento de glicerol, lo que representa el 19 por ciento del peso corporal total del animal, proporción enorme. Esta cantidad de anticongelante permite que los insectos se sobreenfríen hasta −38 grados centígrados.

Puesto que varios animales terrestres pueden evitar con éxito la congelación mediante un fuerte sobreenfriamiento, quizá parezca raro que otros se hayan hecho tolerantes a la congelación, enfrentándose a la tarea mucho más difícil de regular y sobrevivir a la congelación de los líquidos corporales. Pero evitar la congelación tiene sus riesgos. El estado sobreenfriado es metaestable, y la probabilidad de nucleación espontánea por debajo del punto de congelación aumenta a medida que el período de enfriamiento se prolonga y la temperatura continúa bajando. Un enfriamiento por debajo del punto de sobreenfriamiento o el contacto con nucleadores (por ejemplo, como resultado de una herida en la piel) produce una congelación instantánea y fatal. Muchos animales que toleran la congelación pueden haber optado por renunciar a la naturaleza probabilística del sobreenfriamiento en favor de un enfriamiento lento y controlado que, si se hace adecuadamente, permite sobrevivir sin dificultad.

En cierta medida, la razón básica para “preferir” la congelación al sobreenfriamiento se ha perdido en la historia evolutiva de cada especie. Las orugas de la polilla de las agallas que estudiamos comparten su hogar invernal en los tallos de vara de oro con larvas de la mosca de las agallas (*Eurosta solidaginis*). Las orugas se sobreenfrían y evitan la congelación, pero las larvas de mosca de las agallas se congelan cuando la temperatura cae por debajo de unos −8 grados C. Ambos insectos experimentan las mismas condiciones climáticas invernales, pero emplean estrategias de supervivencia opuestas.

No tenemos una respuesta adecuada a esta dicotomía, sólo la observación de un hecho manifiesto: una especie ha desarrollado mecanismos para evitar la congelación, la otra para tolerarla. Las orugas de la polilla de las agallas consiguieron eliminar los nucleadores internos de su cuerpo, tejieron un capullo impermeable para evitar que el hielo ambiental pudiera desencadenar la congelación y después perfeccionaron la estrategia de sobreenfriamiento. En cambio, las larvas de la mosca de las agallas no pueden bloquear, o no bloquean, la acción de nucleadores y, en su lugar, han perfeccionado los sistemas para tolerar la congelación.

¿De qué modo los animales como las larvas de la mosca de las agallas sobreviven, pues, a la congelación? Ya señalamos antes que los cristales de hielo pueden causar extensas lesiones físicas en la estructura interna de las células, así como a las conexiones con mayor grado de organización entre células o a la integridad de los capilares. De hecho, la destrucción causada por el hielo en el interior de las células es tan generalizada que ni siquiera animales tolerantes a la congelación sobreviven al hielo intracelular. Lo mismo puede decirse, a efectos prácticos, de todo tipo de células y tejidos de mamíferos que han sido criopreservados con éxito hasta la fecha. En consecuencia, la tolerancia a la congelación en la naturaleza implica una tolerancia al crecimiento del hielo en los espacios intracelulares llenos de fluido, junto con mecanismos que mantienen líquido el citoplasma.

Para sobrevivir a la congelación, los animales deben emplear adaptaciones bioquímicas específicas que satisfagan tres condiciones básicas. La

primera de ellas es que la formación de hielo debe ser controlada. El crecimiento del hielo ha de iniciarse en los fluidos extracelulares (por ejemplo, el plasma sanguíneo, el líquido abdominal y la orina), de tal manera que el ritmo de congelación se mantenga lento y el tamaño de los cristales pequeño. Para conseguir estos objetivos, los animales tolerantes a la congelación añaden agentes nucleadores específicos a sus fluidos extracelulares. Al proporcionar lugares de enlace que organizan las moléculas de agua en una estructura reticular de hielo, los agentes nucleadores estimulan la cristalización y permiten que ocurra más fácilmente.

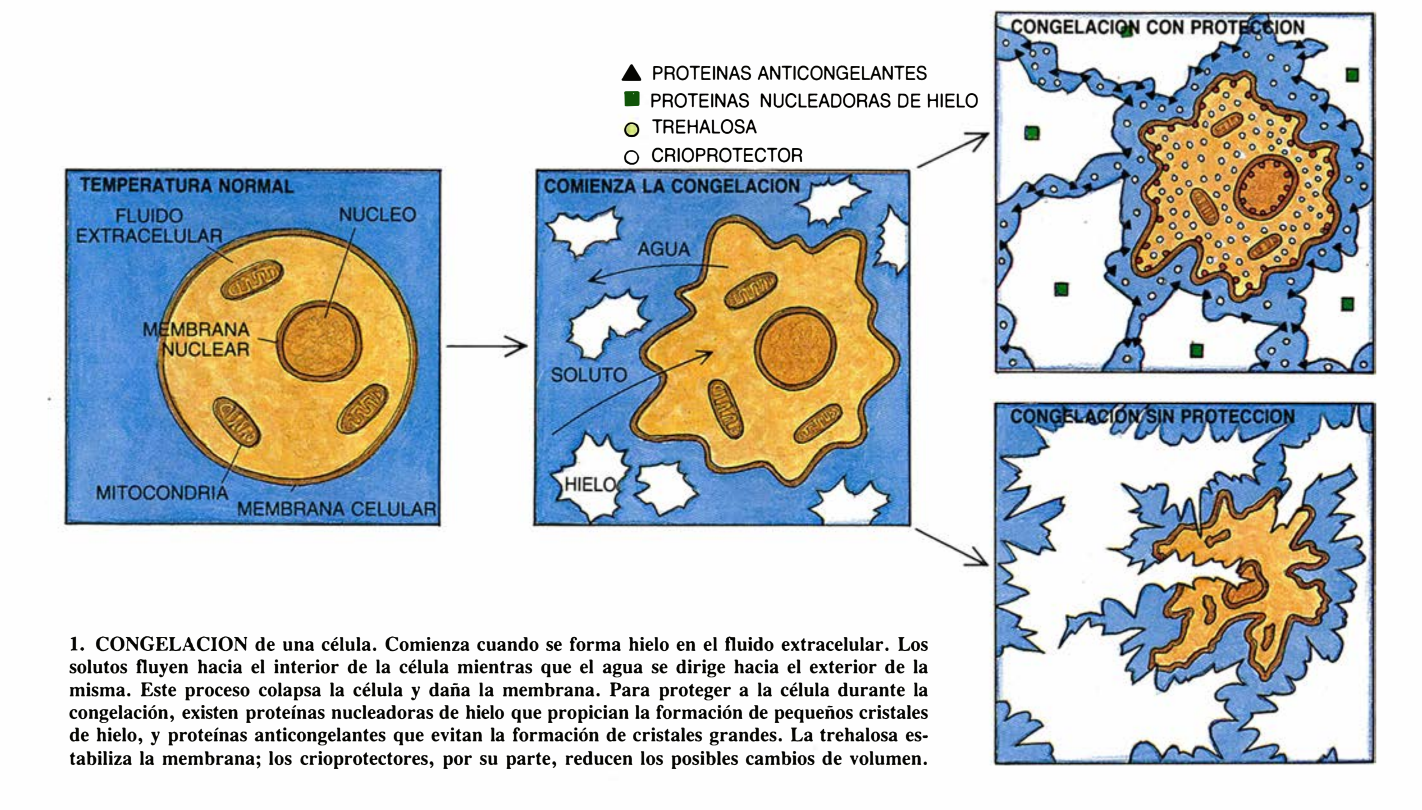
Los nucleadores biológicos en los animales tolerantes a la congelación son, la mayoría de las veces, proteínas sanguíneas específicas (proteínas nucleadoras de hielo), que se sintetizan durante los meses de otoño. La regulación de su producción proviene probablemente de los mismos tipos de señales fotoperiódicas y de la estimulación hormonal, que controlan la síntesis de proteínas anticongelantes en los insectos que evitan la congelación. Las proteínas nucleadoras de hielo provocan la formación de hielo, iniciando en general la cristalización a una temperatura inferior a los dos grados C por debajo del punto de congelación de los fluidos corporales.

Tal proceso minimiza el alcance del sobreenfriamiento; la congelación se convierte así en un episodio bastante lento y controlado que concede a las células disponer de mucho tiempo para ajustarse física y metabólicamente durante la transición al estado congelado. Nuestros estudios de proteínas nucleadoras de hielo en la sangre de ranas de bosque (realizados en colaboración con Jan P. Wolanczyk y John G. Baust, de la Universidad estatal de Nueva York en Binghamton) han demostrado que estas proteínas son muy potentes. Tan sólo un 0,5 por ciento en volumen de sangre de rana añadida a plasma humano es capaz de incrementar la temperatura de nucleación del plasma en siete grados centígrados.

La acción de las proteínas nucleadoras de hielo asegura que el proceso de congelación inicial provoque la dispersión de miles de pequeños cristales de hielo por todos los espacios extracelulares del animal. Sin embargo, los cristalitos de hielo, termodinámicamente inestables, tienden a reconstituirse, con el tiempo, en cristales cada vez mayores, de forma muy semejante a la aparición de cristales de hielo de tamaño apreciable en los helados que se guardan durante mucho tiempo después de abrirlos. Esa recristalización podría ocasionarles daños físicos a los animales, sobre todo en la luz de los capilares y demás espacios delicados; por tanto, los que toleran la congelación precisan un mecanismo que controle el tamaño de los cristales de hielo.

John G. Duman y sus colaboradores, de la Universidad de Notre Dame, identificaron el mecanismo en cuestión. Observaron la intrigante presencia simultánea de proteínas nucleadoras de hielo y anticongelantes que, aparentemente, realizan funciones opuestas en insectos tolerantes a la congelación. Los experimentos no tardaron en mostrar, sin embargo, que las acciones moleculares que permitían que las proteínas anticongelantes bloquearan el crecimiento de cristales de hielo embrionarios eran igualmente eficaces en el bloqueo de la recristalización de los cristales ya existentes. Ambas proteínas, pues, controlan conjuntamente la estructura del hielo: las proteínas nucleadoras de hielo inician la formación de hielo extracelular y las anticongelantes estabilizan los cristales de hielo en un tamaño pequeño e inofensivo.

La segunda condición para sobrevivir a la congelación implica la protección de la estructura y la función celulares. La membrana celular semipermeable, que separa los compartimientos extracelular e intracelular, permite el libre trasiego de agua y de algunos solutos, pero restringe el movimiento de otros compuestos. De modo que cuando se forma hielo fue-



1. CONGELACION de una célula. Comienza cuando se forma hielo en el fluido extracelular. Los solutos fluyen hacia el interior de la célula mientras que el agua se dirige hacia el exterior de la misma. Este proceso colapsa la célula y daña la membrana. Para proteger a la célula durante la congelación, existen proteínas nucleadoras de hielo que propician la formación de pequeños cristales de hielo, y proteínas anticongelantes que evitan la formación de cristales grandes. La trehalosa estabiliza la membrana; los crioprotectores, por su parte, reducen los posibles cambios de volumen.

ra de las células, esto cambia de inmediato el equilibrio de agua y de solutos en el interior de aquéllas.

El hielo es un cristal de agua pura. A medida que el hielo extracelular se va formando, excluye de su estructura solutos tales como sales, azúcares y proteínas. Así, el fluido extracelular restante se hace cada vez más concentrado. Este proceso supone un estrés osmótico para la célula, porque la concentración total de solutos a ambos lados de la membrana celular debe equilibrarse siempre. En respuesta a dicha tensión, del interior de las células fluye agua y los solutos penetran en su interior. El proceso se detiene cuando la concentración de solutos evita, con el tamaño adquirido, la pérdida ulterior de agua en forma de hielo.

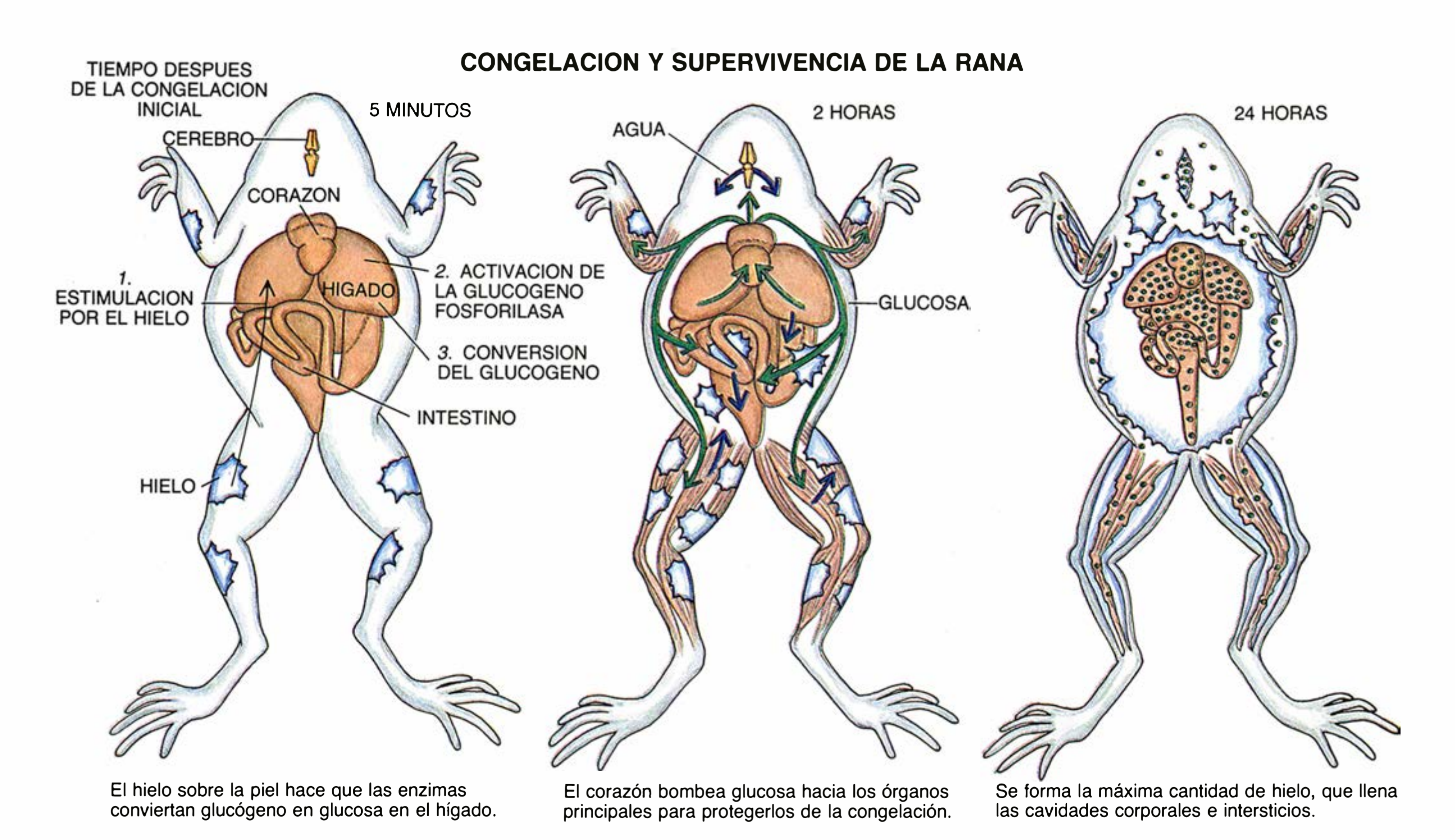
La lesión más grave que cierne su amenaza durante la congelación afecta a la membrana celular. La salida de agua celular provocada por la formación de hielo extracelular reduce rápidamente el volumen celular; la membrana celular se deshincha. Si el volumen celular se reduce por debajo de un mínimo crítico, entonces la bicapa de fosfolípidos de la membrana se comprime tanto que su estructura se rompe. No pueden mantenerse las funciones de transporte de la membrana y se vierte el contenido celular por las roturas de ésta, creándose una puerta de entrada para que el hielo se propague por el interior de la célula. La mayoría de animales resistentes a la congelación alcanzan el volumen celular mínimo crítico cuando alrededor del 65 por ciento del agua corporal total se halla secuestrada en forma de hielo.

Para contrarrestar tales tensiones sobre la estructura celular, los animales tolerantes a la congelación emplean a la vez crioprotectores de membrana y coligativos: compuestos de bajo peso molecular que evitan de maneras diversas las lesiones que se producirían por cambios importantes en el volumen celular durante la congelación. Los crioprotectores de membrana interactúan con los fosfolípidos de la membrana para extender la bicapa y estabilizar la estructura de la membrana a medida que el volumen celular se reduce. Se sabe que la trehalosa, un disacárido, y la prolina, un aminoácido, son los compuestos naturales que realizan esta función. No es sorprendente que los animales tolerantes a la congelación, como las larvas de la mosca de las agallas, acumulen cantidades sustanciales de ambos compuestos durante los meses de otoño previos a su primera exposición a temperaturas de congelación.

Los crioprotectores coligativos ayudan a limitar, mediante la acción osmótica, la cantidad de hielo que puede formarse, el agua que las células pueden perder y, por tanto, la constricción del volumen celular durante la congelación. Cuanto mayor es la concentración de solutos en un fluido, menos hielo se forma a una determinada temperatura y más se puede hacer descender la temperatura de un animal antes de que se alcance el letal 65 por ciento de contenido en hielo. Por tanto, los animales que toleran la congelación añaden elevadas concentraciones de solutos no tóxicos a sus fluidos corporales, y así, cuando se produzca la congelación extracelular, la consiguiente reducción del volumen celular pueda minimizarse.

Con esta finalidad, los insectos tolerantes a la congelación emplean los mismos polihidroxialcoholes que usan las especies que evitan la congelación como protección anticongelante. Las larvas de moscas de las agallas acumulan una enorme reserva de carbohidratos en su cuerpo graso (el equivalente al hígado en los insectos) durante las últimas semanas de alimentación estival. Durante los meses de otoño este glucógeno almacenado, que supone alrededor del 8 al 12 por ciento del peso total del cuerpo de las larvas, se convierte en dos polihidroxialcoholes: glicerol y sorbitol.

Las enzimas implicadas de una manera decisiva en la síntesis de estos compuestos responden tan sólo a las bajas temperaturas. Mientras que la actividad de la mayoría de enzimas y



El hielo sobre la piel hace que las enzimas conviertan glucógeno en glucosa en el hígado.

El corazón bombea glucosa hacia los órganos principales para protegerlos de la congelación.

Se forma la máxima cantidad de hielo, que llena las cavidades corporales e intersticios.

otros procesos metabólicos disminuye con el descenso de la temperatura, temperaturas entre cero y cinco grados C aumentan la actividad de la glucógeno fosforilasa, al estimular la conversión de su forma inactiva a su forma activa (la enzima secciona del glucógeno unidades del azúcar hexosa para iniciar la síntesis). Además, las temperaturas bajas inactivan otras enzimas; en virtud de ello, el flujo de carbono se desvía, desde las rutas normales del catabolismo de los carbohidratos (utilizados para producir energía celular) hacia rutas especiales que conducen a la síntesis de crioprotectores. Los crioprotectores persisten a lo largo de todo el invierno y, después, cuando empieza la primavera, se convierten de nuevo en azúcares para abastecer el desarrollo continuado de los insectos a través de los estadios de pupa y adulto.

El glicerol, el sorbitol y los compuestos relacionados representan, en términos bioquímicos, elecciones excelentes de crioprotectores. Estos compuestos, además de proporcionar las acciones osmóticas necesarias para la regulación del volumen celular durante la congelación, siguen siendo no tóxicos para las células, incluso a concentraciones muy elevadas. No cristalizan espontáneamente a partir de soluciones acuosas a baja temperatura y atraviesan las membranas con entera libertad. Además, esos polihidroxialcoholes estabilizan la estructura de proteínas y enzimas y las protegen de los efectos desnaturalizadores de las temperaturas bajas o de la congelación.





2. RANITA ARBOREA GRIS (*izquierda*)**: se congela bajo el manto de nieve invernal (—8 grados centígrados). Sus pigmentos dérmicos se tornan azules cuando el animal se congela** (*derecha*)**.**

Nuestros estudios de ranas tolerantes a la congelación han revelado un sistema distinto del empleado por los insectos. Las ranas recurren a un crioprotector diferente y siguen un método insólito para desencadenar su síntesis. Las ranas de bosque, las ranitas arbóreas y las ranitas listadas acumulan cantidades masivas de glucosa, el azúcar habitual en la sangre de los vertebrados, durante los episodios de congelación (las ranitas arbóreas grises emplean glicerol). Mientras que la sangre humana tiene un contenido normal de glucosa de unos 50 a 100 miligramos por cada 100 mililitros (en los diabéticos, los niveles de glucosa pueden triplicarse o cuadruplicarse), las ranas de bosque después de congelarse tienen niveles de glucosa en sangre que en promedio suponen 4500 miligramos por cada 100 mililitros. Todos los órganos del cuerpo de la rana contienen a su vez glucosa en concentraciones que parecen óptimas para la protección de cada órgano.

Pero las ranas no acumulan gradualmente sus depósitos de crioprotectores a lo largo de los meses de otoño, como hacen los insectos. Antes bien, esperan hasta que empieza la verdadera congelación en la superficie de su piel. El hielo sobre la piel desencadena una respuesta hormonal o nerviosa que activa, de forma instantánea, la degradación del glucógeno en el hígado, vertiéndose así grandes cantidades de glucosa a la sangre. Hemos detectado niveles crecientes de glucosa en sangre a los cinco minutos de la formación inicial de hielo, y los órganos quedan bien repletos de glucosa en menos de ocho horas, mucho antes de las 24 horas que se precisan aproximadamente para alcanzar los niveles máximos de hielo en el cuerpo que permiten la supervivencia. La rápida síntesis de crioprotector durante la congelación, y la reconversión igualmente rápida a glucógeno hepático cuando las ranas se descongelan, puede ser la clave para evitar los diversos efectos negativos de niveles de glucosa elevados y sostenidos que en el hombre se hallan asociados, por ejemplo, con la diabetes o el proceso de envejecimiento.

Pero si niveles elevados de glucosa pueden ser perjudiciales, ¿por qué utilizan las ranas el azúcar para la crioprotección? Una razón reside en que la glucosa puede producirse rápidamente a partir del glucógeno hepático. La síntesis de crioprotectores en las ranas parece ser una exageración extrema de la respuesta de "luchar o huir", mediada por la adrenalina, que tiene lugar en todos los vertebrados y que incrementa rápidamente la glucosa en sangre durante situaciones de tensión. Así, compuestos que bloquean la acción de la adrenalina en el hígado, como el propranolol, bloquean también la síntesis de glucosa por parte del hígado de la rana durante la congelación.

Sin embargo, nuestros estudios con fibras cardíacas y células hepáticas de la rana de bosque sugieren una razón más crucial para la elección de la glucosa: el azúcar tiene efectos benéficos específicos para la criopreservación de los órganos de los vertebrados. Por ejemplo, fibras ventriculares recobraron su capacidad de contraerse tras la descongelación si se habían congelado en presencia de niveles elevados de glucosa, pero no si se congelaron en una concentración equivalente de glicerol. Puesto que la glucosa y el glicerol proporcionarían los mismos efectos osmóticos para controlar la reducción del volumen celular durante la congelación, la superioridad de la glucosa debe proceder de otras acciones específicas que coadyuven a la supervivencia celular.

Una de estas acciones puede ser el empleo de glucosa como combustible para proporcionar energía en forma de trifosfato de adenosina (ATP) a las células que no tienen acceso al oxígeno aportado por la sangre mientras permanecen congeladas. Además, hemos observado que niveles elevados de glucosa (pero no de glicerol) pueden retardar la biosíntesis de urea en el hígado de la rana, fenómeno que sugiere la posibilidad de que niveles elevados de glucosa intervengan también en la suspensión del metabolismo de los órganos congelados, limitando así los requerimientos energéticos celulares y, con ello, favoreciendo la supervivencia a largo plazo.

La tercera y última condición para sobrevivir a la congelación descansa en el mantenimiento de la viabilidad celular. Aunque la baja temperatura corporal de los animales congelados reduce automáticamente

su tasa metabólica, las células de los animales tolerantes a la congelación deben poseer una capacidad bien desarrollada para sobrevivir sin oxígeno, sin acceso a los combustibles que aporta la sangre y sin ser envenenados por la acumulación de los productos finales del metabolismo que normalmente son eliminados a través de la sangre. El cerebro humano puede resistir quizá tres minutos de flujo sanguíneo interrumpido antes de que empiece la necrosis hística; el riñón o corazón que se extirpan para su trasplante resisten de 6 a 12 horas si se mantienen rodeados de hielo. Pero en nuestro laboratorio las ranas de bosque revivían después de una o dos semanas de congelación.

Estos animales deben continuar generando energía celular en ausencia de oxígeno. Encontramos que las larvas congeladas de mosca de las agallas no presentaban en absoluto alteraciones de sus niveles celulares de ATP cuando eran congeladas durante una semana; la energética en los órganos de la rana permanecía estable durante al menos tres días de congelación. Aun cuando los niveles de ATP bajaron durante una congelación prolongada, lo mismo las larvas de mosca de las agallas que las ranas restauraban rápidamente los niveles de energía celular después de la congelación. De lo que se deduce que los animales tolerantes a la congelación sobreviven sin oxígeno durante períodos prolongados. Poseen mecanismos para generar suficiente ATP a partir de la fermentación de glucógeno o de glucosa, y todos los órganos toleran bastante bien los períodos prolongados de niveles bajos en energía.

La detención metabólica que ocurre durante la congelación puede asimismo revestir gran importancia en la recuperación subsiguiente. La capacidad de reducir mucho la tasa metabólica, con frecuencia a valores bajísimos (del 1 al 10 por ciento de la tasa de reposo normal), es una estrategia adaptativa clave que multitud de animales emplean para sobrevivir a las condiciones ambientales extremadas. Una mengua de la tasa metabólica hasta la décima parte, por ejemplo, supone para los animales multiplicar por diez el tiempo que un determinado depósito de combustibles corporales puede sostener la vida. El ejemplo más familiar es la hibernación de los mamíferos: al entrar en un estado de letargo y disminuir la temperatura corporal hasta cerca de cero grados C, los micromamíferos ahorran hasta el 88 por ciento de la energía que gastarían para sobrevivir al invierno. Numerosos insectos pasan el invierno en diapausa (un estado de desarrollo detenido) y las tortugas que hibernan en el fondo de estanques reducen su tasa metabólica para sobrevivir todo el invierno sin respirar. En las especies que toleran la congelación, pues, la capacidad de reducir su tasa metabólica mientras están congeladas aumenta grandemente las perspectivas de supervivencia a largo plazo.

Las estrategias adaptativas que los animales emplean para sobrevivir a la congelacion son maravillosas en sí mismas. Pero siempre se nos pregunta sobre la aplicación de nuestros estudios a la criopreservación médica, en particular, para órganos susceptibles de trasplante. La primera criopreservación con éxito tuvo lugar en 1949, año en que se hicieron revivir espermatozoides tras haberlos congelado en una solución de glicerol. Desde entonces, se han desarrollado técnicas para congelar suspensiones de muchos tipos de células (espermatozoides, glóbulos sanguíneos rojos y blancos, plaquetas) y tejidos simples (embrionarios, piel, córnea, islotes pancreáticos).

Sin embargo, para los tejidos complejos siguen existiendo problemas físicos. Los investigadores no poseen todavía las técnicas adecuadas para restaurar un órgano funcional después de la congelación. Entre los obstáculos que se presentan en la congelación de órganos están la dificultad de enfriar o caldear homogéneamente una masa orgánica grande para evitar el daño físico por parte del hielo y el problema de instilar o extraer grandes cantidades de crioprotectores de células que no están adaptadas naturalmente a soportar bien una presión osmótica importante. Hay, además, problemas de índole metabólica. Los crioprotectores artificiales (como el dimetil sulfóxido, ampliamente usado) ofrecen una protección física excelente, pero presentan efectos tóxicos sobre el metabolismo celular. Además, la degeneración metabólica ocurre a los pocos minutos de haber separado un órgano de su suministro de oxígeno sanguíneo, y el proceso de enfriamiento daña todavía más los órganos de los mamíferos, que no están diseñados para funcionar a temperaturas muy por debajo de 37 grados C.

Pero las lesiones causadas por la congelación y los principios mediante los que se pueden evitar son los mismos en la criopreservación que en la tolerancia natural a la congelación; algunas respuestas son también idénticas. Por ejemplo, el glicerol y otros alcoholes y carbohidratos de bajo peso molecular se emplean por lo común para la criopreservación médica gracias a su relativa inocuidad y capacidad para moverse con rapidez a través de membranas. Los investigadores también reducen la temperatura por etapas, para disparar la nucleación extracelular a una temperatura inferior a cero grados C que sea benigna, y para evitar la cristalización espontánea en un citoplasma sobreenfriado. En algunos de los estudios más recientes en este campo se está investigando si estabilizadores de membrana o inhibidores metabólicos (que imponen la detención metabólica al inhibir procesos que utilizan ATP) pueden mejorar la supervivencia durante la congelación.

Nuestros estudios sugieren enfoques adicionales. Podrían desarrollarse compuestos sintéticos que imitaran los efectos de las proteínas nucleadoras de hielo o anticongelantes para regular de manera más efectiva la formación de hielo extracelular; una cuidadosa elección del crioprotector (como la glucosa) podría ofrecer, a la vez, defensa física y metabólica durante la congelación. Las investigaciones deben examinar, asimismo, otras estrategias de detención metabólica para preservar la viabilidad de un órgano mientras se halla congelado. Estamos sometiendo a comprobación estas ideas mediante comparación de la supervivencia, tras la congelación, en tejidos de rana de bosque y de rata. Ya sabemos algunos secretos de la tolerancia a la congelación; esperamos continuar sacando otros a la luz.

BIBLIOGRAFIA COMPLEMENTARIA

BIOPHYSICS AND BIOCHEMISTRY AT LOW TEMPERATURES. Felix Franks. Cambridge University Press, 1985.

LIFE IN THE COLD: AN INTRODUCTION TO WINTER ECOLOGY. Peter J. Marchand. University Press of New England, 1987.

THE BIOPHYSICS OF ORGAN CRYOPRESERVATION. Dirigido por David E. Pegg y Armand M. Karow, Jr. Plenum Press, 1988.

FREEZE TOLERANCE IN ANIMALS. Kenneth B. Storey y Janet M. Storey en *Physiological Reviews*, vol. 68, n.° 1, págs. 27-84; 1988.

LIFE IN A FROZEN STATE: ADAPTIVE STRATEGIES FOR NATURAL FREEZE TOLERANCE IN AMPHIBIANS AND REPTILES. Kenneth B. Storey en *American Journal of Physiology*, vol. 258, n.° 3, parte 2, págs. R559-568; 1990.

¿Está en Africa nuestro origen?

La teoría de que todos los humanos descienden de un antepasado africano reciente fue promovida por genéticos que estudiaban poblaciones actuales. El registro fósil corrobora ese modelo

Christopher B. Stringer

En la década de los setenta, una serie de extraordinarios descubrimientos antropológicos en el este africano transformó la interpretación académica de las primeras fases de la evolución humana. Se ganó muy pronto el interés del gran público. Los medios de comunicación siguieron paso a paso los espectaculares hallazgos de antiquísimos restos de homínidos: fósiles keniatas atribuidos a las primeras especies humanas *Homo habilis* y *Homo erectus*, huellas plantares de Tanzania datadas en 3,7 millones de años y huesos de «Lucy» y sus congéneres australopitecos descubiertos en Etiopía.

La misma década asistió al comienzo de otra revolución en la paleontología humana, cuya importancia no se había reconocido debidamente hasta ahora. Esta pacífica revolución tiene que ver con la última fase de los orígenes del hombre —la aparición de gente como nosotros, nuestra especie, *Homo sapiens*, con su amplia variedad de físico y de color.

En el núcleo del problema reside la cuestión de saber cuán especiales son los hombres modernos. ¿Se desarrollaron nuestros cráneos con arcos superciliares pequeños y nuestro grácil esqueleto en circunstancias singulares dentro de una población única o fueron culminación de antiguas tendencias evolutivas que llevaban inexorablemente a la aparición de los hombres modernos por todo el mundo? La controversia probablemente continuará varios años más, mientras las distintas escuelas se lancen unas contra otras el último hallazgo. El debate acabará por centrarse sobre las pautas evolutivas de las variedades geográficas —comúnmente llamadas razas— de *Homo sapiens*. Tres teorías principales pueden distinguirse: modelo multirregional, de dispersión y de flujo génico.

CHISTOPHER B. STRINGER es director del grupo dedicado a los orígenes del hombre del Museo de Historia Natural de Londres, donde viene trabajando desde 1973. Se graduó en antropología por el University College de la capital británica y se doctoró con una tesis sobre homínidos pleistocenos en la Universidad de Bristol. Sus investigaciones se centran en las relaciones evolutivas entre los neandertales y los hombres modernos; también ha publicado estudios de fósiles de *Homo habilis* y *Homo erectus*. Ha realizado campañas en Inglaterra, Gales, Gibraltar y Turquía.

El modelo multirregional sostiene que los caracteres raciales evolucionaron, a través de largos períodos de tiempo, en las mismas regiones donde se encuentran en la actualidad, si exceptuamos los casos de migraciones relativamente recientes. Tal salvedad ocurrió en Norteamérica, donde los amerindios, inuit, europeos y africanos aportaron sus caracteres distintivos durante los últimos 20.000 años. El modelo multirregional, propuesto hace unos años por un grupo de paleoantropólogos, está encabezado en la actualidad por Milford H. Wolpoff, de la Universidad de Michigan en Ann Arbor; lo apoyan algunos genéticos, como James N. Spuhler, del Laboratorio Nacional de Los Alamos. Si esta explicación es correcta, los caracteres considerados raciales deberían ser antiquísimos. Por ejemplo, las grandes narices de los europeos y los robustos pómulos de los aborígenes australianos tendrían que haber derivado, respectivamente, de los neandertales que vivieron en Europa al menos durante un período de 200.000 años, y se extinguieron hace alrededor de unos 30.000 años, y del Hombre de Java, variedad de *Homo erectus* que vivió en esta isla indonesia hace unos 700.000 años.

El modelo de "dispersión desde Africa" propone que un grupo de hombres modernos, racialmente indiferenciados, evolucionaron en Africa y se extendieron por todo el mundo; desarrollaron los rasgos raciales en ese proceso. A tenor de esta hipótesis, defendida por un grupo de genéticos dirigidos por Allan C. Wilson, de la Universidad de California en Berkeley, los antiguos linajes, como los neandertales europeos y el Hombre de Solo del este de Asia (un probable descendiente local del Hombre de Java, que vivió hace unos 100.000 años), no se mezclaron, o se mezclaron muy poco, con los primeros humanos modernos; éstos sustituyeron a aquéllos.

El modelo de flujo génico, o hibridación, retrotrae las poblaciones modernas a una red de antiguos linajes cuya contribución genética variaría de una región a otra y cuya tasa de intercambio aumentaría quizás a medida que fueran evolucionando los hombres modernos. Partidarios de esta interpretación, en los que entran Erik Trinkaus, de la Universidad de Nuevo México, y Fred H. Smith, de la Universidad del Norte de Illinois, opinan que los datos genéticos y paleontológicos actuales son todavía escasos para poder deshacer la madeja tan compleja del origen.

Aunque todos los modelos predicen que las características modernas evolucionaron tarde y que los caracteres raciales surgieron localmente, el desacuerdo reside sobre todo en el ritmo temporal de la diferenciación de las razas. En los modelos multirregional y de flujo génico, los caracteres raciales preceden a la aparición de los modernos; el modelo africano invierte ese orden.

1. LAS RAZAS constituyen variaciones regionales cuyos orígenes deben hallar explicación en el marco de las teorías de la evolución de los hombres modernos. En esta ilustración de los años cuarenta de nuestro siglo, las razas son atribuidas a los hijos de Noé: Jam, Sem y Jafet.

Recientes descubrimientos parecen haber levantado el tono del debate sin resolverlo, si bien ciertos puntos ya parecen aclarados y otros alimentan la esperanza de un rápido progreso. Resaltaré las esperanzas y las dificultades que entrañarían la interpretación de las dos fuentes de pruebas, paleontológicas y genéticas, en lo que concierne a la aparición de los hombres modernos. Los paleontólogos estudian las pruebas externas de huesos y dientes del registro fósil; los genéticos se apoyan en pruebas internas de las células vivas. En mi opinión, el orden presentado en el modelo africano aporta la explicación más coherente de los fenómenos observados por ambos grupos de investigadores. Empecemos por la contribución de los genéticos.

La genética comenzó estudiando la historia de los genes utilizando los rasgos (color de los ojos) o las proteínas (grupos sanguíneos) que los genes codifican; pero el progreso de la técnica posibilitó el análisis del código en el mismo ADN. En las células animales, el ADN se localiza en el núcleo (ADNn), donde controla la mayoría de los aspectos de la estructura orgánica o fenotipo, y en las mitocondrias, diminutos orgánulos que convierten la glucosa en una forma de energía más fácilmente utilizable. El ADN mitocondrial (ADNmt) podría ser un vestigio de una fase de vida libre en la historia del orgánulo, fase que se supone precedió a la simbiosis que dio origen a las células modernas.

Para nuestros propósitos, la mayor diferencia entre los dos tipos de ADN reside en su forma de herencia. Los genes nucleares se heredan de ambos, del padre y la madre, y se combinan al azar en la reproducción sexual; ésta es la razón por la que los hijos no son exactamente idénticos a ninguno de los progenitores. Los genes mitocondriales, por contra, son transmitidos sólo por la madre. Las alteraciones del ADN pueden atribuirse a errores de copia, o mutaciones, mutaciones que ocurren en el ADNmt con mayor frecuencia que en el ADNn. Merced a ese comportamiento, resulta más fácil seguir el rastro de los genes mitocondriales a través de las generaciones y usarlos de reloj molecular (aunque sólo muestran parte de la historia genética del organismo). Tal reloj podría asignar una fecha para cada punto de divergencia en el árbol genealógico.

Para reconocerles fiabilidad a los relojes moleculares deben verificarse dos supuestos: las semejanzas genéticas de los organismos han de ser función de su parentesco y directamente proporcionales al tiempo de divergencia desde un antepasado común. Ambas condiciones requieren que el material genético en estudio mute a una tasa constante y que las mutaciones no sean ni preteridas ni favorecidas por la selección natural. Preterirlas supondría dejar un gen en su condición original, de tal forma que el antecesor común parecería mucho más reciente de lo que en reali-



dad es. La selección de determinadas mutaciones aceleraría el cambio, haciendo que la divergencia parezca mucho más antigua.

Para nuestra fortuna, algunos fragmentos de ADNmt satisfacen tales exigencias, porque, siendo parte del ADN desechable, no parecen afectar al fenotipo. Un postrer requisito para el desarrollo de un reloj molecular válido es que los linajes no se mezclen apenas después de su divergencia. De otro modo, darían la impresión de haber divergido más tarde de lo que lo hicieron.

Se construye un reloj molecular comparando las diferencias genéticas entre las distintas especies o variedades dentro de la misma especie y expresando su parentesco en un árbol. Se calibra luego el árbol, es decir, se "data", por comparación con otro grupo cuya fecha de separación de dicho árbol se conoce ya. La calibración permite a los estudiosos asignar fechas a otras divergencias en el árbol. Importa recordar que tales árboles no son necesariamente verdaderas genealogías; antes bien, configuran agrupaciones que reflejan parecidos y diferencias entre las poblaciones en cuestión. Las pautas evolutivas no se demuestran con los árboles, sino que se infieren a partir de ellos.

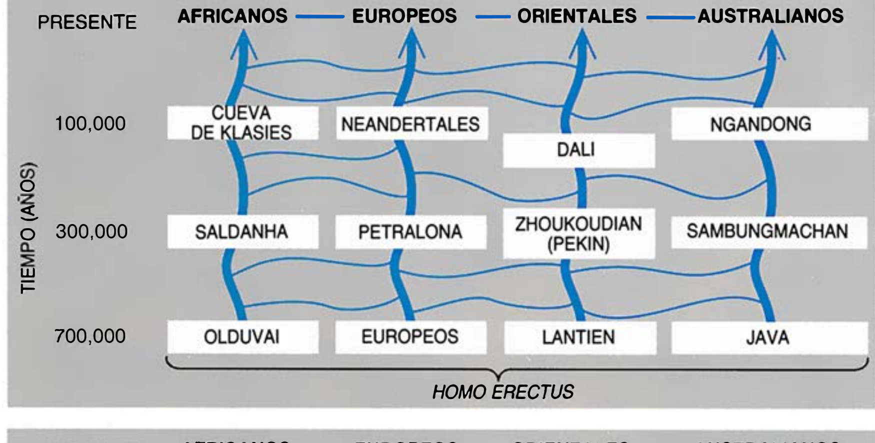
Varios árboles genéticos que ilustran el parentesco de los hombres actuales manifiestan dos unidades fundamentales: los africanos subsaharianos y el resto de las poblaciones humanas. La calibración de la evolución genética a partir de tales árboles puede ser problemática (algunos dirían imposible). Algunos especialistas afirman que todas las diferencias genéticas encontradas entre los seres humanos actuales podrían haberse desarrollado en 100.000 o 200.000 años; otros argumentan que hace falta mucho más tiempo, y se apoyan en que el flujo génico entre poblaciones habría hecho parecer sus diferencias más antiguas de lo que realmente son.

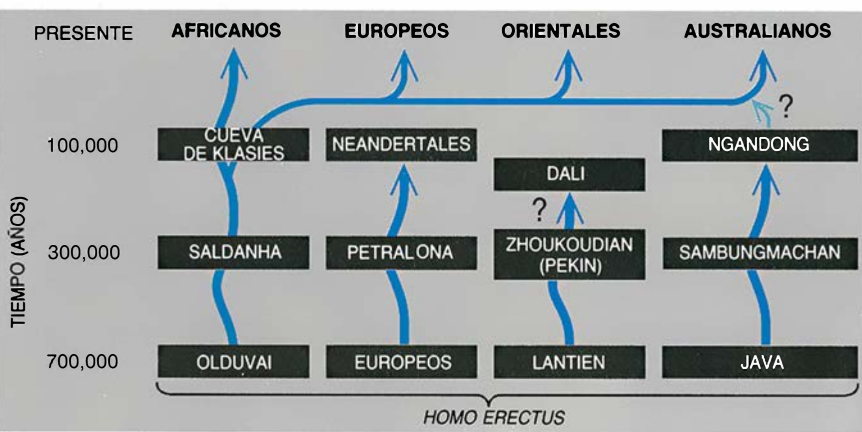
Tan difícil de localizar el antecesor del árbol en el espacio puede serlo como en el tiempo. Un estudio de ADNmt, por ejemplo, podría haberse interpretado en apoyo de un origen asiático lo mismo que africano de los hombres modernos. Si uno prefiere el continente cuyas poblaciones portan los tipos de ADN más próximos a la media general, entonces Asia resulta ser el lugar de origen privilegiado. En este caso, la notable diferencia genética de los koisianos (o bosquimanos) tendría que atribuirse a una acelerada tasa local de mutación.

Por otro lado, si uno parte de la hipótesis de que el ADNmt muta con una tasa constante en todas las poblaciones humanas, Africa adquiere entonces la primacía, y cobra sentido la mayor diferenciación que ofrecen las poblaciones africanas. Conclusión a la que se llegó también desde un estudio subsiguiente del ADNmt, que utilizó el chimpancé como grupo externo a las raíces del hombre en un árbol más comprehensivo y más dilatado en el tiempo. Este es, pues, el primer punto en apoyo de la monogénesis africana.

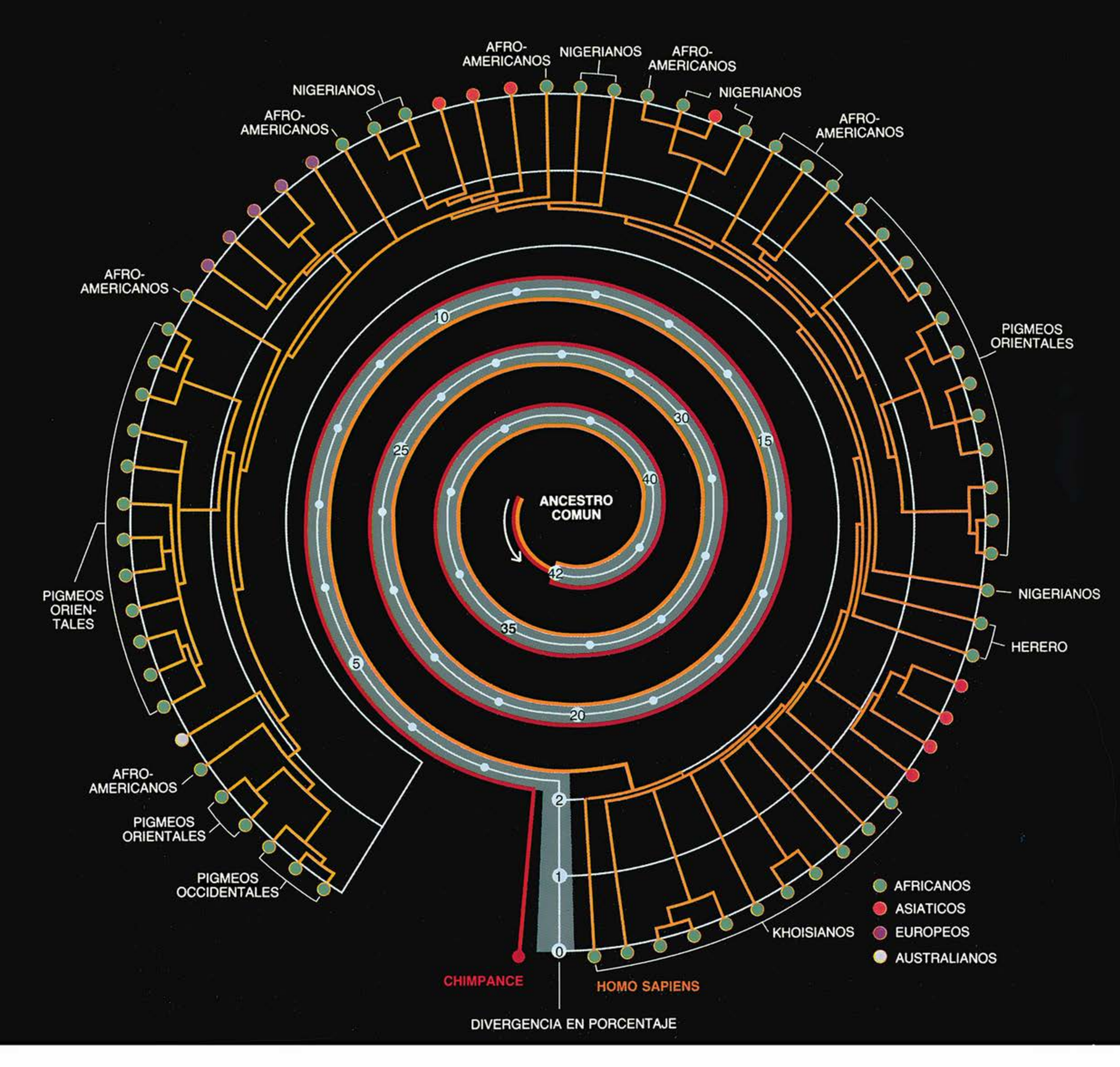
Desacuerdos y simples malentendidos han marcado la controversia sobre la tasa de mutación del ADNmt. También aquí, una razonable resolución del tema favorece la teoría monogenética. Algunos investigadores han argumentado que la tasa promedio de cambio, basada en muy distintos organismos, está alrededor del 0,7 por ciento (es decir, unas 115 mutaciones en la molécula de ADNmt) por cada millón de años. Otros han estimado un cambio del 2 al 4 por ciento (entre 330 y 660 mutaciones) por millón de años. No siempre se tiene en cuenta que los dos cálculos no están midiendo la misma cosa. La tasa más baja mide el cambio a lo largo de un linaje dado; la más rápida mide cambios acumulativos entre di-

3. MAPA DE DISTANCIAS GENETICAS entre poblaciones. Para su confección se atiende a las mutaciones que se han acumulado en un segmento de ADN mitocondrial en rápida evolución. El diagrama abarca una muestra mundial de poblaciones humanas. Como referencia se incluye una muestra de chimpancé *(línea roja)*. La mayor divergencia intrahumanos, 2 por ciento, aparece entre los khoisianos y otros africanos. Esta medición puede visualizarse siguiendo la línea amarilla que va desde el círculo verde más próximo hasta el centro, que representa una muestra de khoisianos, hasta el nodo que une a éste con otras ramas y comparando la distancia con el índice *(barra gris que marea las "seis" del reloj)*. Las líneas del hombre y el chimpancé se separan de un antepasado común hace alrededor de 5 millones de años, tiempo durante el cual han divergido en un 42 por ciento *(centro)*. El gráfico se funda en un estudio dirigido por Linda Vigilant, de la Universidad de California en Berkeley.





2. EL MODELO MULTIRREGIONAL *(arriba)* defiende que los caracteres locales fueron heredados de antiguos homínidos cuyos restos han sido encontrados en yacimientos de Africa, Europa, este de Asia y Australasia. Este modelo atribuye la unidad de la especie moderna *(Homo sapiens)* a un inferido flujo genético entre poblaciones contemporáneas *(líneas horizontales)*. El modelo de monogénesis *(ilustración inferior)* mantiene que las variaciones regionales surgieron más recientemente, después de que los humanos modernos originados en Africa reemplazaran a otros homínidos sin mezclarse con ellos, salvo, quizás, en Australasia. El modelo de flujo de genes (hibridación) se parece al modelo multirregional, pero sitúa el flujo genético principal en estados evolutivos más recientes.



ferentes linajes. La última tasa debería, por tanto, ser doble de la primera. Además, ciertas partes de la molécula de ADNmt parecen evolucionar más deprisa que otras, y las tasas más rápidas de cambio se calculan principalmente en estos fragmentos. Este tipo de consideraciones pueden dar cuenta de muchas de las discrepancias en las cifras.

Tales variaciones en las constantes de cambio pueden provocar una diferencia de varios cientos de miles de años en la edad estimada para el antepasado común de la humanidad. Una tasa lenta o errática situaría nuestro antepasado en hace más de 500.000 años, en el tiempo de *Homo erectus*; otra más rápida, más constante, situaría al antecesor dentro de los últimos 150.000 años, cuando el moderno *Homo sapiens* existía ya probablemente.

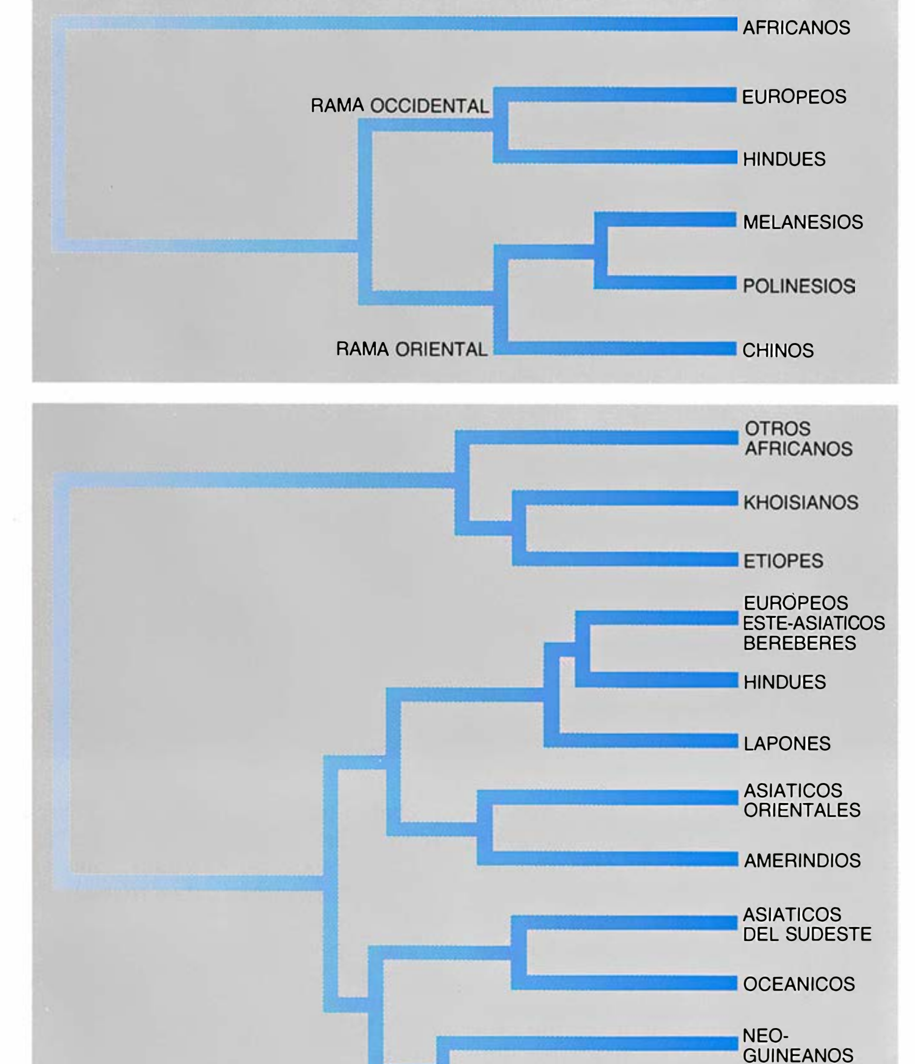
La fecha más reciente parece preferible, pues tiene mayor sentido evolutivo. Si los hombres modernos hubieran comenzado a diferenciarse en los tiempos de *Homo erectus,* algunos de sus subgrupos habrían tenido que desarrollar caracteres esqueléticos modernos independientemente unos de otros. Parece, por contra, más probable que los árboles de ADNmt reflejen la evolución dentro de un acervo génico de humanos, que ya compartían muchas características de aspecto moderno.

Las reconstrucciones genéticas basadas en ADN nuclear implican que los europeos modernos, asiáticos y australasiáticos guarden una relación mutua más estrecha que con los africanos subsaharianos. Resultado difícil de reconciliar con el registro fósil de homínidos de alrededor de 300.000 años, el cual evidencia una ligazón entre las poblaciones de Europa y Africa por un lado y las de China e

Indonesia por otra. Además, la diferenciación progresó de tal suerte que, hace unos 100.000 años, las poblaciones homínidas divergían de una región a otra: había neandertales en Europa y en el oeste asiático, humanos modernos en Africa y oeste asiático, homínidos arcaicos de afinidades inciertas en China y gentes del tipo *Homo erectus* en Java. Sabemos ahora que las poblaciones actuales de Europa, Asia y Australia son genéticamente muy próximas y que las primeras poblaciones modernas que habitaron estas regiones, hace entre 20.000 y 30.000 años, se parecían también mucho unas a otras.

Mi opinión es que estas observaciones avalan la hipótesis monogenética: los patrones demográficos actuales comenzaron, más probablemente, con la dispersión desde Africa de los primeros humanos modernos en los últimos 100.000 años.

La evaluación que hice, 10 años atrás, del registro fósil abonaba la hipótesis de la dispersión africana. Pero se la consideró, hasta hace poco, una conclusión sin base experimental puesto que las técnicas de radiocarbono no podían trabajar con materiales de más de 30.000 o 40.000 años. Dificultad que se está desvaneciendo gracias al desarrollo y extensa aplicación de las técnicas de datación de la familia, o serie, del uranio (U-S), termoluminiscencia (TL) y resonancia de espín electrónico (REE). El primer método puede usarse en sedimentos de cueva, como estalagmitas; el segundo, en sedimentos o sílex quemados por antiguos fuegos, y el tercero en distintos materiales, dientes de animales en particular.



4. DENDROGRAMAS DE AFINIDAD basados en la variación de ADN nuclear *(arriba)* y en sus productos *(abajo)*. Estos árboles respaldan la teoría del origen africano del hombre moderno. El árbol superior está adaptado de un trabajo realizado por Jeffrey C. Long, de la Universidad de Nuevo México; el inferior, de un estudio acometido por Luigi L. Cavalli-Sforza, de la de Stanford.

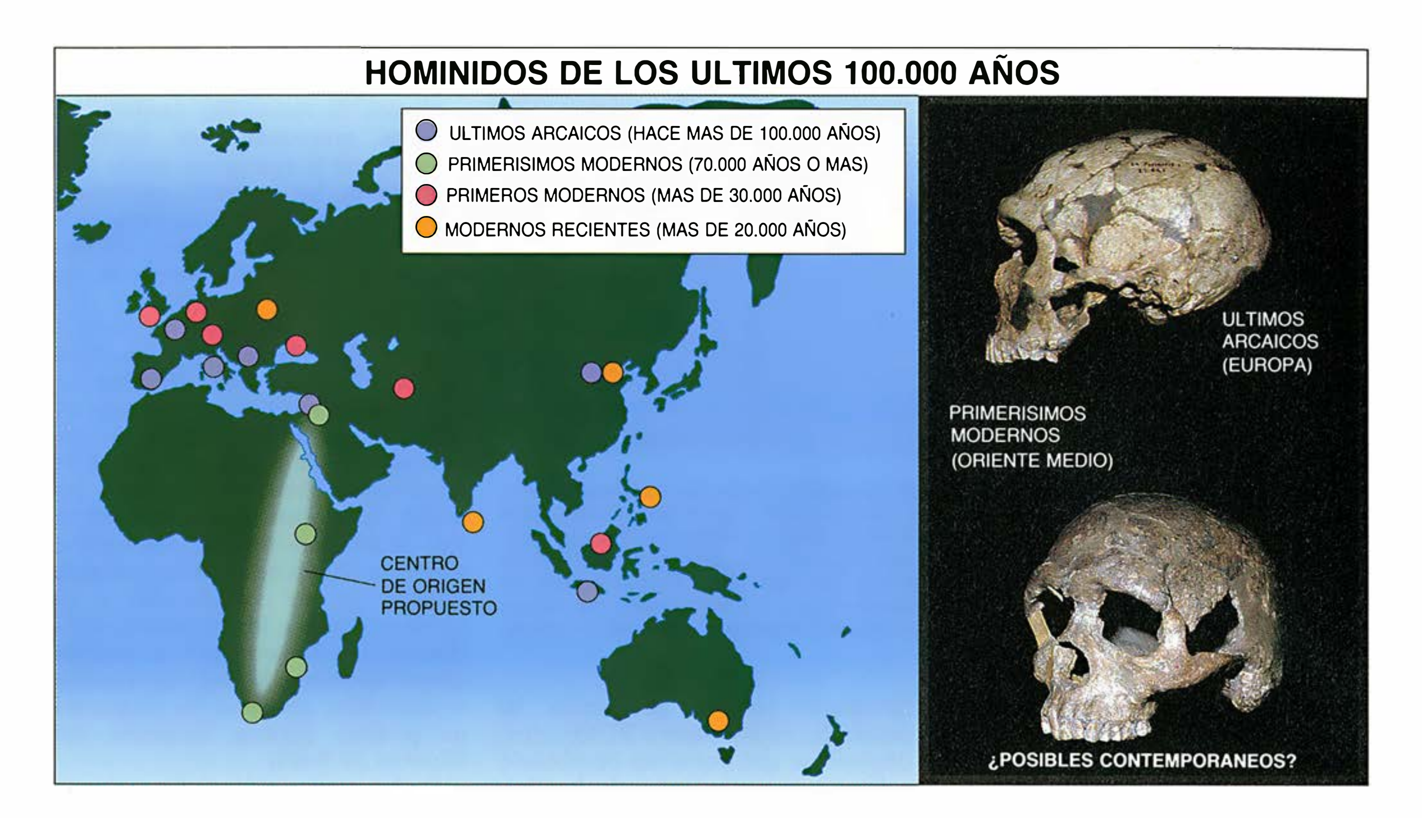
En cada caso, se asigna una fecha a los restos de homínidos mediante la determinación de la edad que muestran los materiales con que aparecen asociados los restos. Aunque las tres técnicas tienen limitaciones, en su conjunto han confirmado que el hombre moderno estaba presente en el Oriente Medio y en Africa cuando los neandertales todavía vivían en Europa y Asia occidental y el hombre de Solo probablemente ocupaba todavía Java.

Se levantó la sospecha de que así pudo suceder a raíz del descubrimiento, hace años, de fragmentos de cráneo, mandíbulas y tres huesos de extremidad con aspecto moderno en las cuevas de la desembocadura del río Klasies y en Border Cave, ambas en Sudáfrica. En esos yacimientos, los restos estaban asociados con artefactos del Paleolítico medio africano (período comprendido entre hace 130.000 y 35.000 años). En Klasies, las fechas U-S y REE situaron hace poco los niveles del Paleolítico medio africano y la mayor parte de los homínidos en más de 90.000 años. En Border Cave, la situación se complica por culpa de la excavación asistemática de los dos primeros homínidos descubiertos. Varios antropólogos han argumentado que tres de los fósiles proceden de niveles que el REE ha fechado en 75.000 años y que el cuarto procede de un nivel hoy datado en unos 60.000 años. Yacimientos de Kenia, Etiopía y Marruecos han proporcionado piezas de huesos anatómicamente modernos, que, en opinión de algunos, podían ser incluso más antiguos que los de Sudáfrica.

Otras pruebas de que los primeros humanos modernos no evolucionaron de las últimas poblaciones primitivas nos las ofrecen recientes reinterpretaciones de los hallazgos de las cuevas de Israel. Vale la pena revisar el criterio sobre estos descubrimientos desde una perspectiva histórica.

En los años treinta, las cuevas de Skhul y Tabun, en el monte Carmelo, proporcionaron esqueletos parciales, resultado probable de enterramientos voluntarios. Se propusieron muchas teorías acerca de las poblaciones representadas en las dos cuevas, pero se fue imponiendo la idea según la cual Tabun había sido habitada por neandertales y Skhul por modernos con facciones primitivas. Entre 1935 y 1975 se excavó una muestra mayor de hombres modernos primitivos en Jebel Qafzeh, cerca de Nazaret; se descubrieron más neandertales en las

HOMINIDOS DE LOS ULTIMOS 100.000 AÑOS



cuevas de Kebara, también en el monte Carmelo y Amud, junto al mar de Galilea. Se aceptó, pues, que en Oriente Medio y en Europa los humanos modernos llegaron después de los neandertales. Podían, en consecuencia, haber descendido de ellos.

La idea se descartó en 1987-88, cuando el empleo de TL y REE permitió datar el neandertal de Kebara en torno a los 60.000 años de antigüedad (tan antiguo como se esperaba), pero la datación de los antiguos humanos modernos de Qafzeh resultó ser unos 40.000 años más remota. Los nuevos cálculos sugirieron que en el Oriente Medio existieron dos linajes humanos. Uno, que partía de antecesores africanos o poblaciones locales mal conocidas y llegaba a la gente cuyos restos fueron encontrados en Qafzeh, dio lugar, diríase, a humanos totalmente modernos. El otro, la línea neandertal, apareció hace unos 70.000 años, quizá como resultado de una migración desde una Europa con creciente glaciarismo. Los neandertales habrían entonces ocupado el área hasta que los humanos modernos la reocuparan en tiempos de la gente de Skhul, hace quizás unos 40.000 años.

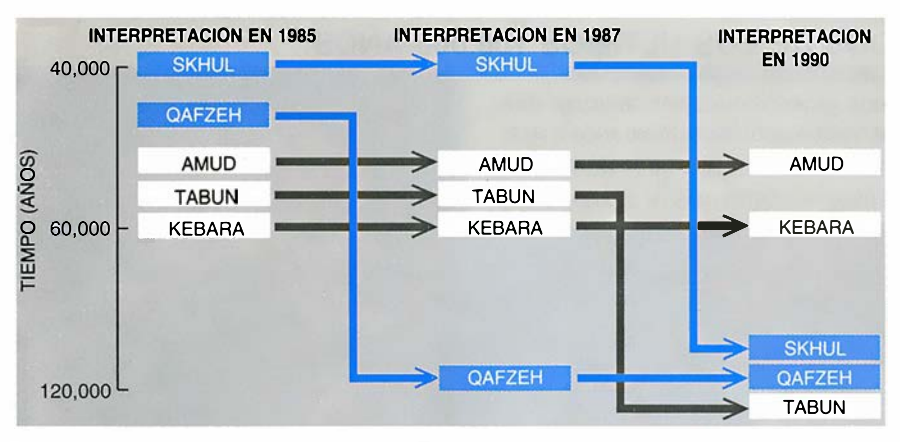
Está por decir la última palabra, ya que las dataciones más recientes de REE, realizadas en 1988-89, fechan dientes de mamífero, procedentes de los niveles de enterramiento humano en Skhul y Tabun, alrededor de los 100.000 y 120.000 años respectivamente. Los nuevos datos parecen confirmar una presencia temprana de humanos modernos, pero añaden la complicación de que también había allí neandertales. Los neandertales (de Tabun) precedieron inmediatamente a los modernos de Qafzeh y Skhul o bien fueron sus aproximados contemporáneos.

Los esquemas cronométricos revisados aportan pruebas importantes para mantener que los humanos modernos constituyen una rama que se separó de los neandertales, una rama que evolucionó en Africa, en el Oriente Medio o en ambas regiones. Además, si los neandertales precedieron y sucedieron también a los antiguos modernos en Oriente Medio, resultaría muy verosímil que los antiguos modernos inmigraran hacia esa área desde otro lugar. Puesto que los registros más antiguos conocidos de los hombres modernos se localizan en Sudáfrica e Israel, y presentan unos 100.000 años de edad, es probable que una población ancestral, más antigua, hubiera vivido en una zona intermedia, quizás en el norte o el este de Africa.

Los nuevos datos apoyan también la opinión de una minoría de investigadores, entre los que me incluyo, para quienes los neandertales habrían constituido una especie distinta (*Homo neanderthalensis*). Otra prueba de una clara separación biológica entre neandertales y modernos viene dada por la persistencia de dos poblaciones cuya identidad se mantuvo separada un dilatado período de tiempo. El neandertal de Kebara podría haber vivido 40.000 años después de que las dos poblaciones hubieran entrado en contacto; sin embargo, este espécimen no muestra signos de hibridación con humanos modernos, hasta el punto de constituir uno de los esqueletos neandertales más robustos y característicos. Por igual razón, los fósiles de primeros modernos de Israel y Líbano, datados entre 30.000 y 40.000 años, no muestran caracteres que pudieran atribuirse a una anterior hibridación con neandertales.

Por tanto, los humanos modernos y los neandertales parecen constituir distintas líneas que divergieron a partir de un antepasado común hace más de 200.000 años: los neandertales evolucionaron en Europa, los modernos en Africa. Los progenitores definitivos de los modernos no pueden ser identificados todavía en el registro fósil, aunque hay quien los ve en materiales encontrados en varias partes de Africa. Pero se trata de muestras que difieren grandemente unas de otras; algunas incluso parecen tan recientes que resulta imposible sean antepasados de los primerísimos humanos modernos.

Los avances recientes de muchos campos convierten a nuestros días en un momento apasionante para el estudio de la aparición del hombre moderno. Se progresa, por ejemplo, en la relación entre el cerebro y sistema vocal con la emergencia de culturas basadas en símbolos y abstracciones; se avanza en la reconstrucción de los lenguajes más antiguos. Cabe que nos encontremos ante las puertas



5. REVOLUCION DE LAS DATACIONES, que ha trastocado el orden que tradicionalmente se había asignado a los restos de los primeros humanos modernos *(azul)* y neandertales *(blanco)*. Hasta hace poco, los paleoantropólogos pensaban que los hombres modernos habían sucedido a los neandertales y podrían haber evolucionado a partir de ellos *(izquierda)*. Los primeros resultados de las técnicas de datación de termoluminiscencia, resonancia de espín electrónico y familias de uranio los situaron, sin embargo, antes y después de los neandertales *(centro)*. La interpretación más reciente publicada de los datos abona la posibilidad de una existencia contemporánea de los dos grupos.

de una teoría unificada que reúna las pruebas fósiles, arqueológicas, genéticas y lingüísticas en apoyo de la hipótesis de monogénesis africana. Quedan, sin embargo, algunas cuestiones que contestar si queremos que esa síntesis globalizadora no se derrumbe como un castillo de naipes. He aquí una muestra:

Las condiciones que existieron sólo en Africa y en ningún otro lugar más deben aclararse para dar cuenta de la evolución de los humanos modernos en aquel continente. Estas condiciones tienen que ver con la ecología y el comportamiento humano, tema del que conocemos muy poco en Africa y en otros muchos lugares.

Debe explicarse la propagación de los primeros humanos modernos desde Africa hacia el oeste de Asia y de ahí al resto del mundo, quizá relacionado este éxodo con cambios climáticos o crecimientos demográficos sumados a una progresiva capacidad adaptativa del hombre. Las causas desencadenantes de la dispersión intervinieron hace de 50.000 a 150.000 años; y ésos u otros factores adicionales persistieron en vigor mientras el hombre moderno continuó abriéndose en un intervalo de más de 10.000 años.

Hay que demostrar las ventajas somáticas, cerebrales y conductuales, asociadas quizá con la aparición del lenguaje moderno, para explicar por qué los modernos reemplazaron a los homínidos que los precedieron. A uno le gustaría saber por qué tales ventajas tardaron 70.000 años en producir su éxito evolutivo. Muchos arqueólogos afirman que no existen pruebas de que los primeros modernos poseyeran ventajas claras; su refinamiento cultural, proclaman, se desarrolló en los últimos 30.000 años, después de que hubieran desplazado a las líneas precedentes de homínidos.

Tiene que demostrarse, o negarse, el cruce entre humanos arcaicos y primeros humanos modernos para resolver los desacuerdos entre los modelos multirregional y de hibridación por un lado y el modelo africano por otro. En la aseveración más radical de la hipótesis africana —sugerida por algunas interpretaciones de las pruebas aportadas por el ADNmt— la sustitución habría sido total. Mas, ¿podría un grupo de cazadores-recolectores reemplazar a otro, en todo un continente, sin el más mínimo intercambio? Algunos, por contra, distinguen signos de hibridación entre los últimos arcaicos y los primeros modernos en áreas tales como el este de Europa. Se afirma también, y en ello hay general acuerdo, que los neandertales del oeste de Europa cambiaron sus técnicas instrumentales establecidas desde hacía tiempo para remedar las empleadas por sus contemporáneos los primeros modernos. ¿Podrían haber aprendido de los recién llegados o incluso haberse mezclado con ellos?

Existen serias dudas sobre la fiabilidad con que pueda reinterpretarse el registro fósil en busca de signos de hibridación. Además, aun cuando se hubiese producido la hibridación, podría haber sido tan local que los genes de los híbridos habrían pasado a las poblaciones modernas en muy bajas frecuencias, si es que pasaban. En este caso, podrían desenterrarse fósiles con signos de hibridación, aun cuando los estudios genéticos modernos no lograran descubrirlos.

Hemos de reconciliar también las pruebas procedentes de Extremo Oriente con la emergente teoría unificadora de los hombres modernos. Los fósiles descubiertos en China a lo largo de los últimos 15 años parecen representar una línea distinta de los contemporáneos neandertales del oeste y de los predecesores locales, como el Hombre de Pekín (de cerca de 350.000 años). Esta línea hipotética podría ofrecer una explicación alternativa a la hipótesis africana o quizás unirse a ella en una teoría dual de antecesores afro-asiáticos. Pero los fósiles de la misma región datados en sólo 20.000 años tienden, por contra, a reforzar la hipótesis exclusivamente africana ya que corresponden a un momento en que los caracteres específicamente asiáticos estaban todavía desarrollándose. Si esta aseveración se consolida, entonces los caracteres no podrían haberse heredado del Hombre de Pekín.

De la comparación de las variaciones dentarias actuales se desprende que los caracteres dentales de las gentes del sureste asiático son los que más se acercan al promedio de la condición humana. Ahora bien, lo mismo que en el caso de los trabajos sobre ADNmt examinados antes, las pruebas dentarias no pueden aducirse como una manifestación de un origen asiático del hombre moderno. Los asiáticos estarían más cerca de la media en muchos aspectos si los humanos se hubieran originado en Africa y se hubieran extendido primero por Asia y después por Europa, Australasia y las Américas.

Australia presenta problemas particularmente difíciles, habida cuenta de que los restos de sus poblaciones más antiguas muestran una notable diversidad morfológica. Parece probable que los primeros en arribar a Australia llegaron hace unos 50.000 años, procedentes de algún lugar del sureste asiático. Se desconoce el tipo físico de los inmigrantes más antiguos. En cualquier caso, 25.000 años atrás unos cráneos australianos eran más gráciles y otros mucho más robustos que la inmensa mayoría de sus equivalentes contemporáneos, los primeros de Europa y Asia. Mueve a perplejidad, sobre todo, la presencia de variedades robustas, porque contradice, diríase, la tendencia global hacia un cráneo más grácil.

La aparición de tan amplias diferencias en el cráneo —incluyendo las variedades robustas— puede atribuirse a un proceso de microevolución local dentro de Australia, desencade-

nado por factores ambientales y de comportamiento. Sin embargo, algunos multirregionalistas, Wollpoff entre ellos, hacen responsable de esa variación a la influencia de las antiguas gentes de Solo (Ngandong), cuyos restos se han encontrado en Java. Desde esta perspectiva, su cruzamiento con antiguas poblaciones modernas dejó vestigios en Australia y, por ende, en los actuales aborígenes australianos. El éxito del modelo del origen africano depende de la resolución que se dé a estas cuestiones.

Los próximos años, con el progreso de las técnicas de investigación, traerán probablemente muchas sorpresas a paleoantropólogos y genéticos. Está ganando terreno la idea de que los hombres modernos existieron en Africa y Oriente Medio antes de que aparecieran en Europa y Australasia. Hay también un acuerdo creciente sobre el patrón de relaciones genéticas entre las poblaciones modernas que favorece un origen africano en última instancia, aunque la fecha de ese origen siga en la arena de la controversia.

Se desenterrarán nuevos fósiles y podría extraerse nueva información de huesos, genes y útiles líticos. En un futuro no muy lejano cabría la posibilidad de estudiar ADN sacado de restos fósiles y compararlo con los patrones ancestrales reconstruidos hoy a partir de las poblaciones vivientes. Estos y otros datos que aún no tenemos de tiempos pasados permitirán un día explicar las discrepancias en los resultados según los presentan los que estudian las poblaciones vivas y los que estudian las extinguidas. Pasos que habrá que dar para reconstruir y entender las vastas migraciones que intervinieron, según todos los indicios, entre el origen de los humanos modernos y su conquista de la Tierra.

BIBLIOGRAFIA COMPLEMENTARIA

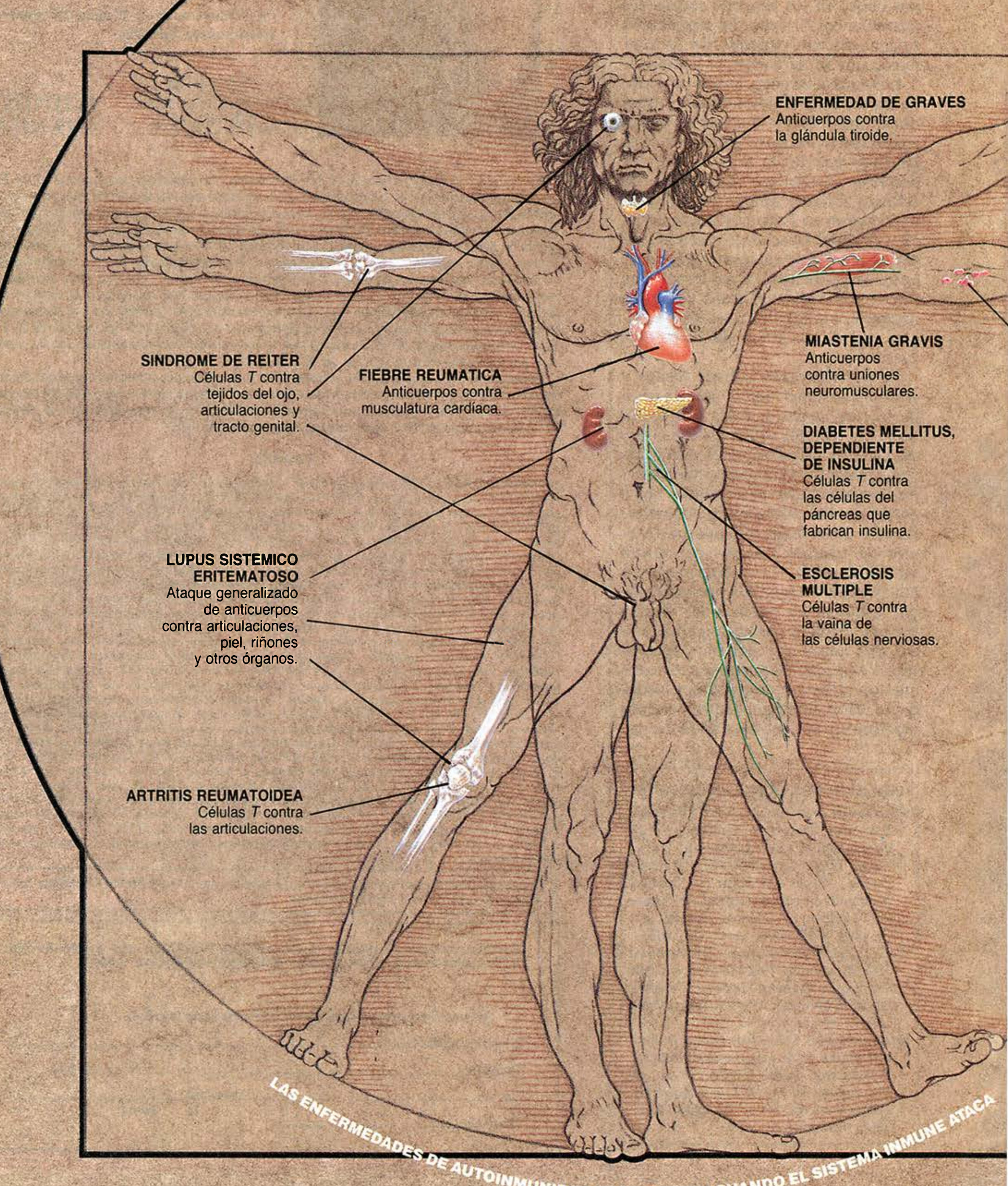
RECONSTRUCTION OF HUMAN EVOLUTION: BRINGING TOGETHER GENERTIC, ARCHAEOLOGICAL, AND LINGUISTIC DATA. Luigi Luca Cavalli-Sforza, Alberto Piazza, Paolo Menozzi y Joanna Mountain en *Proceedings of the National Academy of Sciences*, vol. 85, n.º 16, págs. 6002-6006; agosto de 1988.

THE HUMAN REVOLUTION: BEHAVIORAL AND BIOLOGICAL PERSPECTIVES ON THE ORIGINS OF MODERN HUMANS. Dirigido por Paul Mellars y Chris Stringer. Princeton University Press, 1989.

MITOCHONDRIAL DNA SEQUENCES IN SINGLE HAIRS FROM A SOUTHERN AFRICAN POPULATION. Linda Vigilant et al. en *Proceedings of the National Academy of Sciences*, vol. 86, n.º 23, págs. 9350-9354; diciembre de 1989.

EL CUERPO CONTRA SI MISMO

John Rennie



Hay buenas perspectivas para el tratamiento de los trastornos del sistema autoinmune, pero su funcionamiento aún suscita división de opiniones entre los investigadores.

PSORIASIS
Células *T* contra la piel.

TEJIDOS SANOS

David A. Hafler se encamina hacia el laboratorio para enseñarnos sus células. Vistas al microscopio, las doradas células del fondo de la caja de cultivo tienen un tierno e inocente aspecto. Sin embargo, explica Hafler, reaccionan con contundencia ante la presencia de algunos componentes de la vaina de mielina lipídica que envuelve las neuronas. "Se vuelven locas —asegura— cuando ven las proteínas básicas de la mielina." Hafler habla como un padre orgulloso, y con razón, ya que las células que nos enseña son realmente suyas: proceden de su propia sangre.

La presencia de ese tipo de células puede ser motivo de preocupación. Son linfocitos *T* que atacan a la mielina, y se sospecha que son responsables de la esclerosis múltiple (EM), una enfermedad degenerativa que se produce cuando el sistema inmune del cuerpo ataca al sistema nervioso central. La EM puede llegar a producir ceguera, debilitamiento general o incluso parálisis.

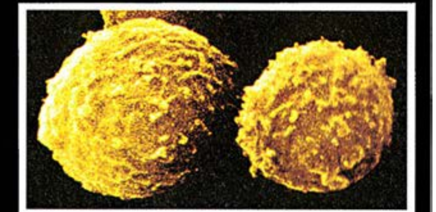
Mas, para Hafler, el hecho de que él y otras personas sanas porten esas células, y no padezcan EM, constituye una buena señal. Con Howard L. Weiner, jefe del programa de EM en el Brigham and Women's Hospital de Boston, donde aquél trabaja, cree que esas células potencialmente peligrosas se mantienen a raya gracias al concurso de unas células supresoras, todavía por identificar, y de otros mecanismos del sistema inmune.

Los investigadores de Boston tratan de utilizar esos mecanismos para curar la EM y otros trastornos de la autoinmunidad, en los que un sistema inmune díscolo se muestra incapaz de distinguir entre lo "propio" (el cuerpo) y lo "ajeno" (los microorganismos invasores). De momento, han vacunado con células *T* inactivadas a un pequeño grupo de pacientes con EM. Confían en que esas inoculaciones desencadenen reacciones supresoras en los pacientes.

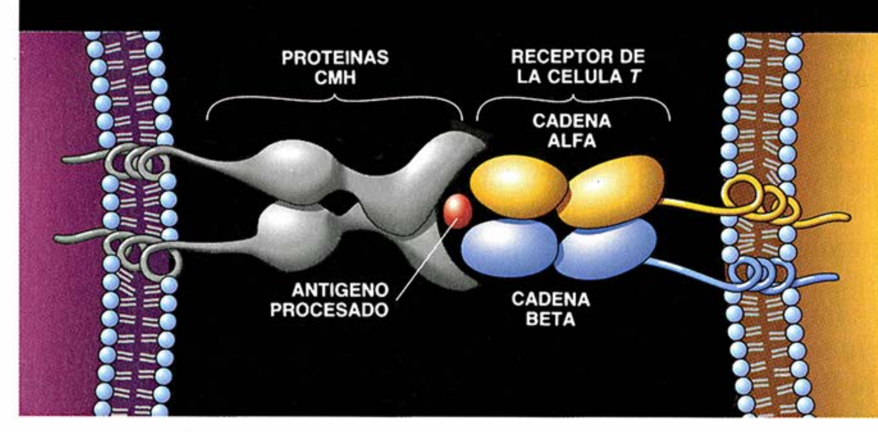
Hafler y Weiner dedican su esfuerzo investigador a la búsqueda de mejores tratamientos contra las enfermedades de autoinmunidad. Otros han comenzado ensayos clínicos con una amplia gama de anticuerpos hechos a medida, que detectan y destruyen las células problemáticas. También se exploran otras terapias, basadas en trasplantes especiales y administración de proteínas por vía oral, pero están aún en fase muy embrionaria.

Amén de la EM, otras muchas enfermedades se consideran de autoinmunidad: la artritis reumatoidea, la diabetes mellitus dependiente de insulina (Tipo I) y el lupus sistémico eritematoso, por nombrar sólo algunas de las más comunes. El tratamiento ideal para esas enfermedades sería el que restaurase





Los linfocitos *T* (*derecha*) destruyen células infectadas y foráneas; normalmente supervisan la mayoría de las reacciones inmunológicas, incluidas muchas de las que causan enfermedades de autoinmunidad. Las células *T* citotóxicas ("asesinas") atacan directamente a sus blancos. Las células *T* coadyuvantes (los "cómplices") producen sustancias que intensifican la respuesta inmune de las células asesinas y otros componentes del sistema inmune. El receptor antigénico de la célula *T* (*abajo*) es la llave de la actividad de esas células. El receptor consta de una cadena alfa de proteína y otra beta. El receptor es caprichoso: sólo se une a un tipo específico de antígeno, que debe ser presentado por proteínas específicas del complejo principal de histocompatibilidad (CMH), sobre la superficie de ciertas células.



de forma segura y permanente la tolerancia del paciente hacia los tejidos acosados. Los actuales presentan serios inconvenientes. Suelen emplearse drogas antiinflamatorias o inmunosupresores generales, como los esteroides y la ciclosporina. La mayoría de esos fármacos, caros, presentan a menudo efectos secundarios. Además, los médicos no se atreven a prescribir remedios tan drásticos a pacientes en las etapas precoces de la autoinmunidad, cuando serían más beneficiosos.

Las nuevas estrategias son muy prometedoras, pero en algunos aspectos los clínicos trabajan a ciegas. Muchos de los tratamientos parten del supuesto de que algunos tipos de células suprimen las células autorreactivas, causantes de la enfermedad autoinmune. Sin embargo, nadie ha podido aún identificar tales células supresoras.

De hecho, la mayoría de los inmunólogos abrigan serias dudas sobre la teoría de las células supresoras. Se inclinan, en cambio, por dos teorías alternativas, que cuentan con el apoyo de un buen soporte experimental: la deleción clonal y la anergia clonal. Según esas teorías, el sistema inmune tolera a las moléculas del propio cuerpo porque elimina a las células autorreactivas (deleción) o las inactiva (anergia).

"Diez años atrás, la mayoría pensaba que todas las formas de autotolerancia respondían a un mecanismo de supresión", comenta Harald von Boehmer, del Instituto de Inmunología de Basilea y pionero en la búsqueda de la demostración de la existencia de la deleción clonal. "Nunca se llegó a identificar convincentemente las células supresoras." Philippa Marrack, del Centro Nacional Judío de Inmunología y Medicina Respiratoria en Denver y defensora de primera hora de la deleción clonal, asiente: "Todos los experimentos que sentaron las bases de la teoría de la supresión se realizaron hace 10 años, y no ha habido aportaciones sustanciales desde entonces."

Sin embargo, los defensores de alguna forma de supresión no se rinden. Aducen recientes datos clínicos y ensayos en animales como pruebas de que la supresión está lejos de ser una idea marchita. Irun R. Cohen, del Instituto Weizmann de Ciencias, paladín de la supresión, insiste en que tanto la deleción como la anergia clonales "dejan sin explicar muchas de las cosas que conocemos". Sostiene que, para entender las enfermedades de autoinmunidad, es crucial entender el comportamiento del sistema inmune como un todo organizado, y no como la actividad bioquímica de células individuales. "Nadie quiere hablar de «comportamiento» mientras se le pueda achacar todo a una molécula, pero lo siento, ésa es la naturaleza del bicho", apostilla.

La controversia recuerda la historia de los tres ciegos que describen un elefante, y cada uno lo pinta de manera distinta. Los inmunólogos que estudian la biología molecular de la autotolerancia llegan a la deleción y la anergia clonales; los que se centran en el comportamiento, descubren la supresión.

Quizá no sea imprescindible conocer todo el tramado teórico para encontrar una terapia contra la autoinmunidad. Algunas ideas, aunque sólo sean parcialmente ciertas, pueden abrir el camino hacia aproximaciones válidas. Señala Judith Kapp, de la Universidad de Washington en St. Louis y partidaria de la supresión: "Así es como trabaja la medicina clínica. No siempre están a la espera de teorías. A veces, desarrollan un tratamiento, y nosotros nos volvemos locos intentando averiguar por qué funciona."

El nacimiento de una idea

Ronald H. Schwartz, jefe del laboratorio de inmunología celular y molecular en el Instituto Nacional de Alergia y Enfermedades Infecciosas (NIAID), describe lo vaivenes que han sufrido durante décadas las teorías que tratan de explicar el reconocimiento entre lo propio y lo ajeno. "Durante los años que van de 1930 a 1950, la gente se hacía las preguntas adecuadas sobre la autotolerancia —explica Schwartz—, pero nunca obtuvieron las respuestas correctas, probablemente porque no entendían las bases del sistema.".

Ya entonces, lo inmunólogos sabían que la sangre y la linfa contenían moléculas proteínicas, los anticuerpos, capaces de reconocer antígenos específicos, o moléculas diana. Sospechaban también que ciertos leucocitos, como linfocitos y macrófagos, constituían una pieza importante en la respuesta inmune, probablemente produciendo anticuerpos e ingiriendo bacterias. Pero nadie sabía con exactitud cómo reconocían a los antígenos las células y los anticuerpos del sistema inmune, y mucho menos cómo distinguían entre propios y extraños.

El punto de inflexión sobrevino en 1955, cuando F. Macfarlane Burnet, del Instituto Australiano de Investigaciones Médicas Walter y Eliza Hall, formuló la hipótesis que con el tiempo le valdría el premio Nobel. Burnet propuso que cada linfocito reaccionaba sólo contra un tipo de antígeno, pero la enorme variedad de linfoci-

tos producida aleatoriamente por el sistema inmune proporcionaba protección contra cualquier antígeno imaginable. Sugirió también que la respuesta inmune dependía de la multiplicación selectiva de clones de linfocitos (genéticamente idénticos). Cuando una célula encontraba a su antígeno, se multiplicaba, aumentando así la capacidad de respuesta frente a dicho antígeno en el futuro.

Si el sistema inmune producía aleatoriamente las células que reaccionan contra los antígenos, inevitablemente fabricaría células contra sus propios antígenos. Burnet señaló, sin embargo, que si existiese un mecanismo para eliminar los clones autorreactivos, éste podría inducir tolerancia a los autoantígenos, abriendo un portillo, estratégicamente situado, en las defensas naturales del cuerpo.

Las grandes virtudes de la hipótesis de la selección clonal de Burnet fueron su poder explicativo y su simplicidad. Estas características cobraron nuevo valor en las décadas siguientes, pues la inmunología pronto se convirtió en un campo necesitado de ideas sencillas y unificadoras. Pero cuanto más penetraba el experimentador en el sistema inmune, más complejo e intrincado parecía.

Los inmunólogos distinguieron dos tipos de linfocitos: *B* y *T*. Las células *B*, que maduran en la médula ósea, reconocen a los antígenos y comienzan a fabricar anticuerpos. Los linfocitos *T*, que maduran en el timo (una glándula situada, en el hombre, encima del corazón), sólo reconocen antígenos situados en la superficie de células infectadas o anormales. Para colmo de desgracias, había al menos dos tipos de células *T*, las citotóxicas, o "asesinas", que son células que atacan a los tejidos infectados, y las células coadyuvantes, encargadas de promover la respuesta inmune. Como la teoría de Burnet ponía un poco de orden en este maremágnum de células, fue bien aceptada.

Sin embargo, las pruebas que confirmasen la idea de Burnet, según la cual la autotolerancia era causada por la deleción de clones de células *T*, tardaban en llegar. Para verificar que se eliminaban las células que reaccionaban contra los autoantígenos, los inmunólogos debían conocer el destino y la afinidad antigénica de células individuales del cuerpo, algo que se escapaba de las posibilidades técnicas de los años cincuenta y sesenta.

A decir verdad, los primeros experimentos realizados para demostrar que la deleción clonal era la causa de la tolerancia fueron desalentadores. Algunos inmunólogos clínicos encontraron anticuerpos contra proteínas normales, como la insulina, en la sangre de personas sanas. Varios grupos de investigadores inyectaron autoantígenos radiactivos en animales y hallaron que algunos anticuerpos y linfocitos reaccionaban contra ellos, señal inequívoca de que las correspondientes células autorreactivas no habían desaparecido. Obviamente, esos hallazgos indicaban que algunas células autorreactivas sobrevivían, al menos ocasionalmente.

"Como la teoría principal había recibido un directo en la mandíbula —comenta gráficamente Schwartz— quedaba abierto el camino para otros puntos de vista sobre la tolerancia. Y así comenzó la era de la supresión." En 1971, Richard K. Gershon, de la Universidad de Yale, propuso la existencia de células *T* supresoras especializadas, cuya misión consistía en patrullar por el cuerpo para neutralizar la respuesta de las células autorreactivas. Se postularon hipotéticas redes de supresión para explicar el delicado equilibrio mutuo entre grandes cantidades de células asesinas, coadyuvantes y supresoras. Todas las respuestas inmunes, ya fuesen contra lo propio o lo ajeno, eran interpretadas como alteraciones del equilibrio celular normal.

La teoría de la supresión no tuvo adversario hasta la revolución biotecnológica de finales de los años setenta y principios de los ochenta. Las teorías sobre la supresión celular sufrieron entonces una derrota más severa que la padecida antes por la deleción clonal; de pronto, la deleción clonal volvió por sus fueros perdidos.

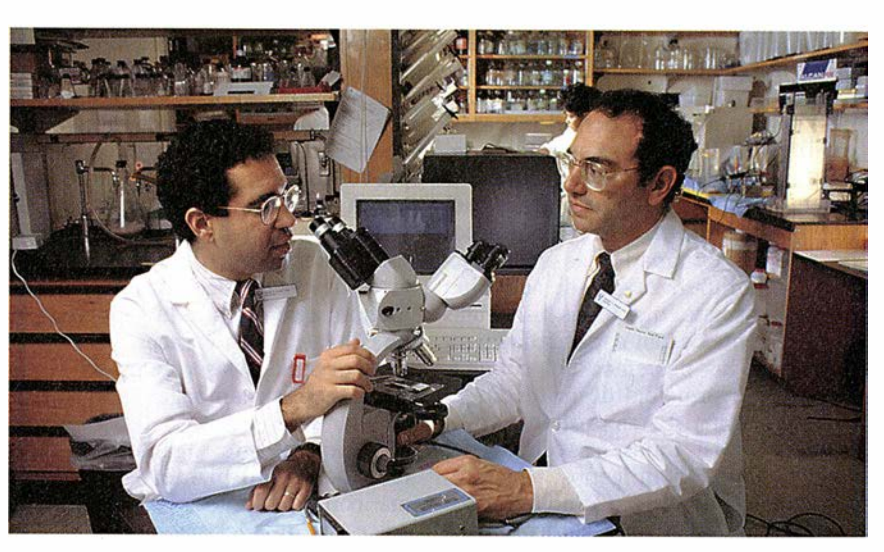
Tres importantes conquistas metodológicas provocaron el cambio. En primer lugar, el descubrimiento de las técnicas para cultivar clones de células *T* en el laboratorio. Más podría aprenderse sobre las interacciones químicas de las células *T* estudiando poblaciones homogéneas, que con mezclas de células *T*. El segundo avance determinante, relacionado con el primero, fue la posibilidad de producir anticuerpos monoclonales. Estas sondas de anticuerpos idénticos podrían localizar proteínas efectoras (diana) en cualquier parte del cuerpo. El tercero fue la técnica del ADN recombinante, que permitió aislar y manipular genes.

Esas técnicas dieron paso a una avalancha de descubrimientos, que aún continúa, relacionados con los mecanismos implicados en el reconocimiento antigénico y la respuesta inmunitaria. Las moléculas receptoras que hacen que las células *T* se comporten de manera específica podían identificarse y analizarse estructural y genéticamente. Por vez primera, la tecnología bioquímica permitía estudiar la autotolerancia molecular.

Sometida al rigor de un examen tan concienzudo, la teoría de la supresión entró en crisis. Gershon y otros habían pronosticado que las células supresoras tendrían moléculas receptoras insólitas y otras características, pero no se pudo encontrar ni rastro de esas moléculas. La existencia de series distintas de células supresoras empezó a perder credibilidad. Sin ellas, las teorías sobre la supresión se resquebrajaron a principio de los ochenta. Nuevos aires soplaban en los laboratorios: lo que no se puede encontrar, afirmábase, es que no existe.

Selección de células

La misma metodología que hundió la teoría de la supresión favoreció los estudios sobre la deleción clonal. Los investigadores podían estudiar el des-



1. INVESTIGAN LA ESCLEROSIS MULTIPLE. David A. Hafler (*izquierda*) y Howard Weiner esperan tratar la autoinmunidad suprimiendo las células del sistema inmune que atacan al cuerpo.



tino de células particulares. Introducían en un ratón, por ejemplo, genes para un receptor específico de célula *T*. En el ratón transgénico que resultaba, todas las células *T* respondían a un único tipo de antígeno, lo que facilitaba su estudio. Una vez más, la deleción fue el foco de atención de los investigadores que estudiaban el problema de la autotolerancia.

En 1988, 33 años después de que Burnet propusiera la idea, aparecieron finalmente las primeras pruebas sólidas de que la deleción clonal se daba en la naturaleza. En estudios independientes, Marrack, von Boehmer y H. Robson MacDonald, del Instituto suizo Ludwig de Investigaciones sobre el Cáncer, demostraron que ciertos clones autorreactivos de células *T* en vías de maduración eran eliminados del timo y no llegaban nunca al resto del cuerpo.

Estudios posteriores de Marrack y otros han aclarado algunos aspectos del proceso que siguen esas deleciones, si bien no describen el mecanismo preciso. Dos etapas selectivas determinan que la multitud de células *T* inmaduras que entran en el timo configuren un repertorio inmunológico, del que son eliminadas las autorreactivas. La primera etapa, de selección positiva, asegura que todas las células supervivientes están capacitadas para reconocer las proteínas del complejo mayor o principal de histocompatibilidad (CMH). Estas proteínas, presentes en casi todas las células del cuerpo, albergan antígenos, y los presentan a las células *T*. Como esas proteínas difieren de unos individuos a otros, resulta imprescindible que las células *T* las reconozcan.

La segunda etapa, de selección negativa, es decisiva para la autotolerancia. Elimina las peligrosas células *T* autorreactivas exponiéndolas en el timo a un batiburrillo de autoantígenos. Las jóvenes células *T* que caen en la trampa, uniéndose agresivamente a los autoantígenos, mueren.

Marrack y otros tratan de resolver ahora el misterio que rodea la muerte de las células *T* durante la selección negativa. Después de todo, reconocer antígenos es lo que se supone que deben hacer las células *T*. Marrack intuye que la clave puede estar en la edad de las células y la fuerza con que se unen a sus antígenos. Las células *T* jóvenes de alta afinidad antigénica pueden morir por sobredosis de estimulación, provocada por las células que les presentan los antígenos.

Un detalle interesante de esta hipótesis de la afinidad es que explica la presencia de las células *T* autorreactivas en la sangre de personas sanas. Dichas células no causan la enfermedad porque no son suficientemente agresivas. Si las afinidades antigénicas de esas células son tan bajas que no provocan una resuesta inmune patológica, el timo no las elimina. Los trastornos de autoinmunidad podrían ser consecuencia de un descuido o error de cálculo del mecanismo de selección, que permite la supervivencia de células *T* más voraces.

Ahora bien, las células de Hafler reconocen, sin ambages, a la mielina en cuanto la ven. Y Marrack es consciente de las interrogantes que dicho reconocimiento plantea. "El problema con la hipótesis de la afinidad –dice– es que debemos preguntarnos si los animales cambian la afinidad sin que sepamos cómo."

Jonathan Sprent, del Instituto de Investigación de la Clínica Scripps de La Jolla, se muestra tajante: "Hablar de afinidades es ganas de hablar por hablar." Los receptores de las células *T* tienen una estructura tan compleja que no es fácil caracterizar sus actividades relativas en términos bioquímicos simples.

Existe, además, un problema fundamental por resolver. Para explicar todo el fenómeno de la tolerancia en función de la teoría de la deleción clonal, es preciso suponer que todos los posibles antígenos del cuerpo se manifiesten en el timo, y así puedan seleccionarse sus correspondientes células *T*. Algunos investigadores piensan que esta idea es verosímil, ya que el timo es punto de encuentro de ma-

crófagos y muchas otras células que presentan antígenos de las que circulan por el cuerpo. Los macrófagos, que van limpiando los tejidos de células muertas, podrían recoger una muestra representativa de autoantígenos. Las moléculas solubles de la sangre y la linfa también circulan por el timo. Sprent, al menos, sí cree que el timo tiene la eficacia suficiente para almacenar los autoantígenos necesarios en una selección eficaz de las células *T*.

Pero la propia Marrack, defensora acérrima de la deleción clonal, reconoce que puede haber puntos débiles en esa teoría. "En algunos casos, podría ser que los autoantígenos no fueran capaces de llegar al timo. Se pueden quedar retenidos en el cerebro, en el dedo gordo o en cualquier otro sitio", dice. "Es preciso desarrollar otro mecanismo que tenga en cuenta esta posibilidad."

La teoría de la anergia clonal surgió como posible apoyo o mecanismo alternativo para explicar la autotolerancia. El término "anergia" fue acuñado por Gustav J. V. Nossal, del Instituto Hall, que describió el fenómeno, a mediados de los años setenta, cuando estudiaba células *B* en cultivo. En determinadas circunstancias, observó, los antígenos no activaban a las células *T*, sino todo lo contrario: inhibían su respuesta. Las células no mueren, sino que permanecen en ese estado de abulia, o anergia, a menos que entren en contacto con grandes cantidades de células *T* coadyuvantes, o interleucina-2, una linfocina o sustancia similar a la hormona del crecimiento que segregan las células *T* coadyuvantes.

Casi una década después, Schwartz y Marc K. Jenkins, que ahora da clases en la Universidad de Minnesota en Minneapolis, descubrieron el mismo fenómeno en células *T* cultivadas. Para explicar sus resultados, Jenkins y Schwartz propusieron un modelo de activación de células *T* basado en la existencia de dos señales.

Activa, inactiva

El modelo de dos señales propone, en definitiva, que la unión de un receptor con un antígeno no basta, por sí sola, para desencadenar una respuesta de las células *T*. Las células deben recibir algún tipo de señal adicional de las presentadoras de antígenos. Si no se recibe esta segunda señal, la celula *T* se anergiza, y permanece así a menos que sea "activada" por las células coadyuvantes.

Sólo ciertos tipos de células especializadas —los macrófagos, por ejemplo— podrían enviar la segunda señal que necesitan las células *T* para activarse y multiplicarse. La mayoría de las células sólo serían capaces de presentar el antígeno. La anergia clonal podría servir así de refuerzo de la deleción en el fenómeno de la autotolerancia: las células *T* autorreactivas que escapan al proceso de deleción en el timo quedarían inactivadas la primera vez que reconociesen un antígeno en una célula somática.

Desde 1987, muchos investigadores han venido sosteniendo que la anergia clonal es un fenómeno que afecta a las células *B* y *T*, no sólo en el tubo de ensayo sino también en el cuerpo del animal.

La anergia no es lo mismo que la supresión, a pesar de algunas semejanzas superficiales en sus efectos. La supresión depende del efecto inhibidor de una célula sobre otra. La anergia es una reacción que la célula sufre cuando no recibe todas las órdenes necesarias para proliferar. Algunos investigadores que trabajan en proyectos relacionados con la supresión argumentan, por contra, que la supresión mediada por células podría servirse de los mecanismos bioquímicos que producen la anergia.

Ocurre que, según quien la exponga, la anergia clonal parece tomar diferentes tonalidades. Schwartz, por ejemplo, cree que las células *T* se anergizan fundamentalmente fuera del timo, en los nódulos linfáticos periféricos. Pero B. J. Fowlkes y Fred Ramsdell, del NIAID, han encontrado, tras estudios con animales, que se produce anergia dentro del timo. En los estudios de Schwartz, con células cultivadas, las que eran anergizadas permanecían en ese estado a menos que fuesen activadas por las células coadyuvantes. Mas los experimentos que Fowlkes y Ramsdell están realizando ahora en ratones sugieren que las células anergizadas acaban activándose por sí solas si dejan de encontrarse con su antígeno específico.

Nadie sabe, no obstante, si la anergia que Fowlkes y Ramsdell han estudiado es la misma que Schwartz y Jenkins han observado en los cultivos celulares. Si en el sistema inmune coexistiesen varias formas distintas de anergia, todos tendrían razón. Sprent, por su lado, tercia displicente: "Tienes una célula, la pones junto a un antígeno y no responde, pero si le añades linfocinas, entonces despierta. Esta es más o menos la definición actual de anergia."

Sprent ha encontrado células que se comportan como si estuviesen anergizadas sin haberse puesto nunca en contacto con su antígeno. Para él, la anergia se convierte en "algo absurdo" si se utiliza la ausencia de res-



¿SUPRESION SUPRIMIDA?

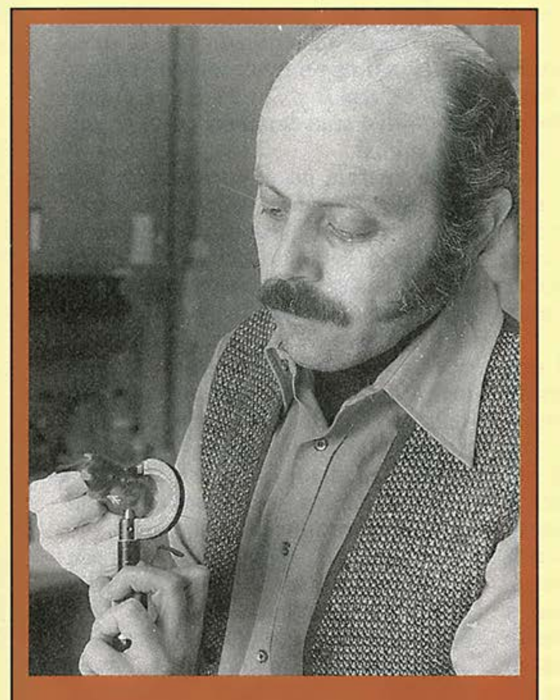
"Hace dos años, ironías del destino, fui invitado a presidir una sesión científica sobre células supresoras, en un momento en el que el 98 por ciento de los inmunólogos no quería ni oír hablar de su posible existencia", recuerda Ronald Germain, del Instituto Nacional de la Alergia y Enfermedades Infecciosas. Y añade: "creo que la gente pensaba que aquello era un castigo bien merecido, por haber trabajado antes en el campo de la supresión".

Sin embargo, Germain empezó la sesión con diapositivas de antiguos artículos relacionados con la supresión, todos ellos escritos por respetados inmunólogos, muchos de los cuales se habían ya distanciado de la teoría. "Le quité el polvo a toda esta historia de la supresión, y recordé que no estábamos hablando de una banda de lunáticos."

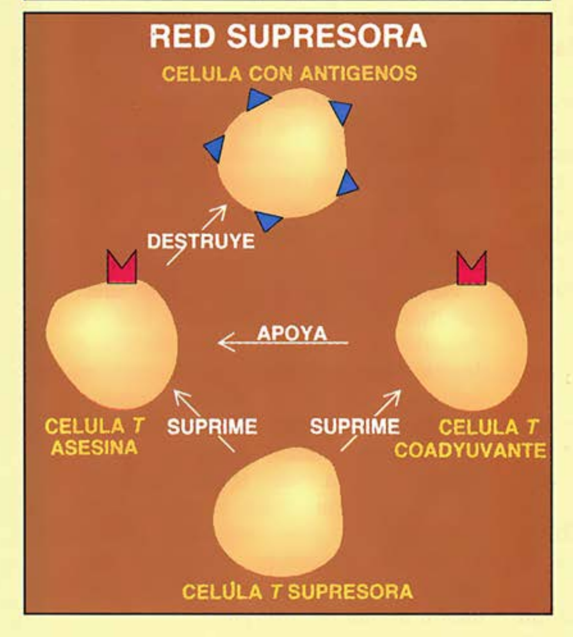
La teoría de la supresión para explicar la inmunotolerancia no goza de buen nombre. Nadie recibe un duro para investigar en ese campo, por el fuerte rechazo que la idea suscita en las comisiones de evaluación. Por eso, algunos prefieren emplear el término "inmunorregulación", más inocuo.

¿Cómo pudo caer en semejante desgracia una teoría que hace diez años ocupaba el centro de las discusiones inmunológicas? "En su época dorada, tenía un vehemente y persuasivo portavoz en la persona de Dick Gershon", comenta Germain. Richard K. Gershon, inmunólogo ya fallecido de la Universidad de Yale, realizó muchos experimentos que sugerían la existencia de células supresoras activas, cuya misión era limitar la actividad de otras células *T*. Durante los años setenta, defendió la idea de que el sistema inmune se suprimía a sí mismo a través de una complicada red de interacciones celulares de ese tipo. Jonathan Sprent, del Instituto de Investigación de la Clínica Scripps, lo recuerda así: "Todos estaban de su parte, pero todo resultaba increíblemente complicado."

En esa complejidad enmarañada residía buena parte de su problema. Fueron contados los laboratorios que crearon sistemas experimentales de supresión que funcionaran, y no había manera de reproducir los experimentos con fiabilidad. A medida que se acumulaban los resultados discrepantes, las redes reguladoras que se proponían eran cada vez más y más "barrocas". Con el tiempo, los investigadores empezaron a cuestionarse la existencia misma de las células supresoras.



Richard K. Gershon propuso la idea, ahora discutida, de que en el sistema inmune existen células supresoras que regulan a otras células.



Comenzó a expandirse una atmósfera de desasosiego. "En cierto sentido (la teoría de la supresión) representaba lo peor de la inmunología: muchos castillos en el aire y experimentos poco afortunados, ninguno de los cuales podía ser confirmado con rigor." Con la muerte de Gershon, en 1983, la teoría de la supresión perdió su más ardiente y elocuente defensor.

Sin una teoría en la que inspirarse, y con fuertes recortes presupuestarios "mucha gente abandonó el campo, desapareció la energía que aportaba la competencia entre diversos laboratorios que trabajaban en la supresión, y los progresos cayeron en picado", comenta Judith Kapp, de la Universidad de Washington.

Pero las teorías sobre la supresión no han muerto. Fuera de los Estados Unidos, aún son tomadas en serio por muchos inmunólogos de prestigio. "Creo que se adapta mejor a la forma de pensar de los europeos. Es más filosófica", dice Irun R. Cohen, que estudia la tolerancia en el Instituto Weizmann de Ciencias, de Israel. Las teorías sobre la supresión inspiran sus trabajos sobre los efectos terapéuticos de la vacunación con células *T* en animales con trastornos de autoinmunidad. "No es atractiva para la gente que está acostumbrada a pensar que el reduccionismo es lo más sagrado de la ciencia", dice Cohen.

Harald von Boehmer, del Instituto de Inmunología de Basilea, tiene sus dudas sobre la supresión, pero confiesa que "hay algunos fenómenos que se explican mejor con ella". Al preguntarle si piensa que sus colegas europeos son más receptivos a la idea de la supresión que los norteamericanos, contesta diplomáticamente: "Algunos norteamericanos se han declarado abiertamente en contra (de ella), pero entre ellos los hay que han cambiado de opinión bastante a menudo en el pasado. Tal vez es sólo que sus opiniones son las más vehementes."

A decir verdad, casi todos los inmunólogos aceptan la existencia de algún complejo fenómeno regulador que guíe el sistema inmune.

"Estoy seguro de que al final se demostrará la existencia de las células supresoras", comenta Sprent. "Más discutible es que se trate de una familia de células diferenciadas. Para mí, constituyen células *T* normales que matan a otras células *T* mediante el reconocimiento de su receptor."

puesta como único criterio. "Estoy seguro de que si se toman algunas células *T* con baja afinidad por un antígeno se comportarían igual."

Estas reservas no preocupan a Fowlkes. "Sabemos que las células están allí, y que no responden, cuando los controles siguen respondiendo", dice. "No cabe sino pensar que las células han sufrido algún tipo de rediferenciación." Ramsdell remacha. "No existe aún una definición molecular de anergia universalmente aceptada —reconoce—, pero es un concepto útil."

Mas es la falta de una definición molecular lo que resta respetabilidad a las teorías de la supresión. Por consiguiente, resultará difícil para quien labore en otros campos entender la preferencia de los inmunólogos por la anergia y su distanciamiento de la hipótesis de la supresión.

"No hay duda de que el fenómeno de la supresión existe", prosigue Hafler. "Se pueden transferir respuestas biológicas con poblaciones de células *T*. De lo que no hemos sido capaces es de aislar una población estable de células supresoras para definirlas molecularmente." La razón por la que nadie ha sido capaz de clonar una célula *T* supresora, sospecha, estriba en el carácter transitorio de la capacidad de supresión. "Nuestros trabajos de clonación de células *T* humanas sugieren que las células pueden transmitir algún tipo de señal supresora en ciertos momentos de sus vidas, pero en otros momentos no", dice Hafler.

A pesar de la desalentadora oposición, algunos investigadores continúan intentando redefinir en términos bioquímicos los efectos supresores que observan. Entre ellos, Kapp, de Saint Louis, gran parte de cuyo trabajo reciente tiene que ver con los mecanismos naturales que crean tolerancia inmunológica a la insulina, una facultad de la que normalmente carecen las personas con diabetes dependiente de insulina. Kapp ha identificado clones de células *T* que reaccionan ante la insulina como si ésta fuese un antígeno, pero también otras del mismo tipo que parecen conferir tolerancia a la insulina, pues su presencia parece disuadir a las primeras de que ataquen a la insulina. Estas presuntas células supresoras no portan moléculas receptoras inusuales, ni se distinguen de las células *T* asesinas por sus marcadores moleculares.

Kapp y sus colegas estudian ahora la tolerancia en ratones transgénicos que fabrican insulina humana. Han establecido ya que los ratones transgénicos toleran la insulina humana, aunque tienen células *T* anti-insulina en sus órganos periféricos. "No hemos establecido el mecanismo de tolerancia para esas células. Lo que es cierto es que no están irreversiblemente anergizadas", comenta. "Y sabemos que la deleción clonal no contribuye mucho, si es que lo hace, a la falta de respuesta."

Kapp cree que la deleción clonal, la anergia clonal y la supresión operan codo con codo en el sistema inmune, induciendo posiblemente la tolerancia a diferentes antígenos. La concentración de un antígeno en el timo puede determinar qué mecanismo actúa: las proteínas unidas a la membrana que se expresan en la mayoría de las células, como las del CMH, podrían tolerarse a causa de la deleción, mientras que la insulina y otras proteínas solubles, que alcanzan sólo concentraciones medianamente bajas en la sangre, tal vez se toleren por supresión.

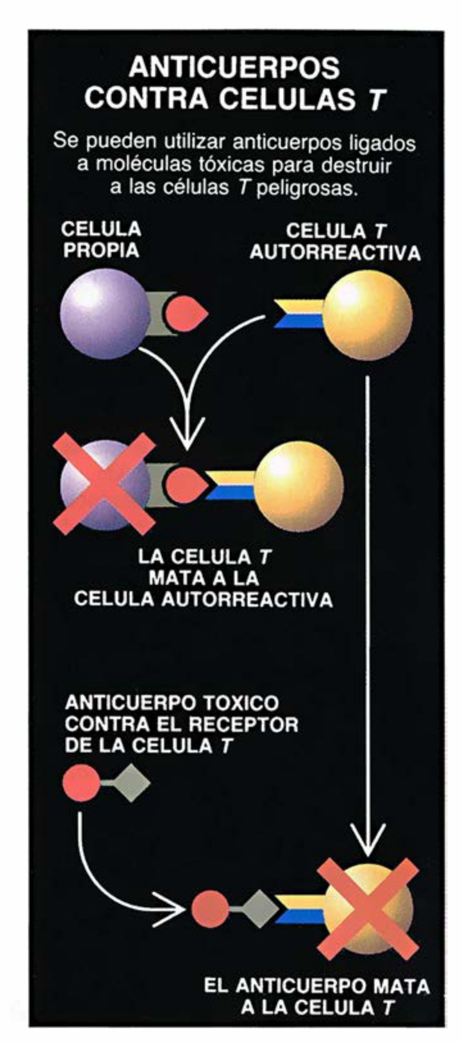
El puente

Si la teoría de Kapp es correcta, aumentando gradualmente las cantidades de insulina expresadas en los ratones transgénicos, "deberíamos alcanzar un punto en el que se produjese deleción clonal. Si lo conseguimos, podremos entender mejor las reglas que gobiernan el proceso", señala ella.

Otros inmunólogos se hacen eco de las reflexiones de Kapp. "No estoy seguro de que [la deleción clonal] explique por qué no reaccionamos contra nuestra propia mielina, o nuestra propia tiroglobulina o nuestro propio receptor de la acetilcolinesterasa", confiesa Lawrence Steinman, de la facultad de medicina de la Universidad de Stanford. "Creo que habrá que tender un puente entre los estudios que se realizan actualmente sobre selección tímica y esas enfermedades de autoinmunidad específicas de los tejidos."

Aunque los inmunólogos coinciden con Steinman en que sus conocimientos sobre la autotolerancia no alcanzan aún a explicar cómo se producen los trastornos de autoinmunidad, algunos investigadores han decidido lanzarse a aplicar los conocimientos obtenidos con el desarrollo de las terapias. Una buena parte de esos trabajos tienen como punto de mira a las células *T* asociadas específicamente con las enfermedades.

A principios del pasado verano, el grupo de Steinman, y Hafler y Weiner, por separado, hicieron público que habían identificado las células *T* responsables de la EM. Los antígenos receptores en esas células parecen contener invariablemente una proteína de un reducido grupo de cuatro, todas muy parecidas, llamadas cadenas beta. (Los receptores de las células *T* son heteroproteínas que constan de una cadena alfa y otra beta, de las que pueden existir muchas formas genéticamente determinadas.)



Si se conocieran las células *T* asociadas específicamente con una enfermedad de autoinmunidad, podrían diseñarse terapias que las eliminasen o suprimiesen selectivamente. Tratamientos que incluirían anticuerpos monoclonales altamente específicos y péptidos a medida que sólo se unen a células *T* activas o a las proteínas CMH de las células presentadoras de antígenos. Para hacer más eficaces esas terapias, los anticuerpos contra las células *T* autorreactivas se ligarían con moléculas tóxicas, formando los llamados inmunoconjugados, que no sólo incapacitarían las células objetivo, sino que las matarían.

Un buen ejemplo de los logros y posibles obstáculos que plantea la hipótesis basada en los anticuerpos monoclonales es el trabajo sobre la encefalomielitis alérgica inducida experimentalmente (EAE), una enfermedad

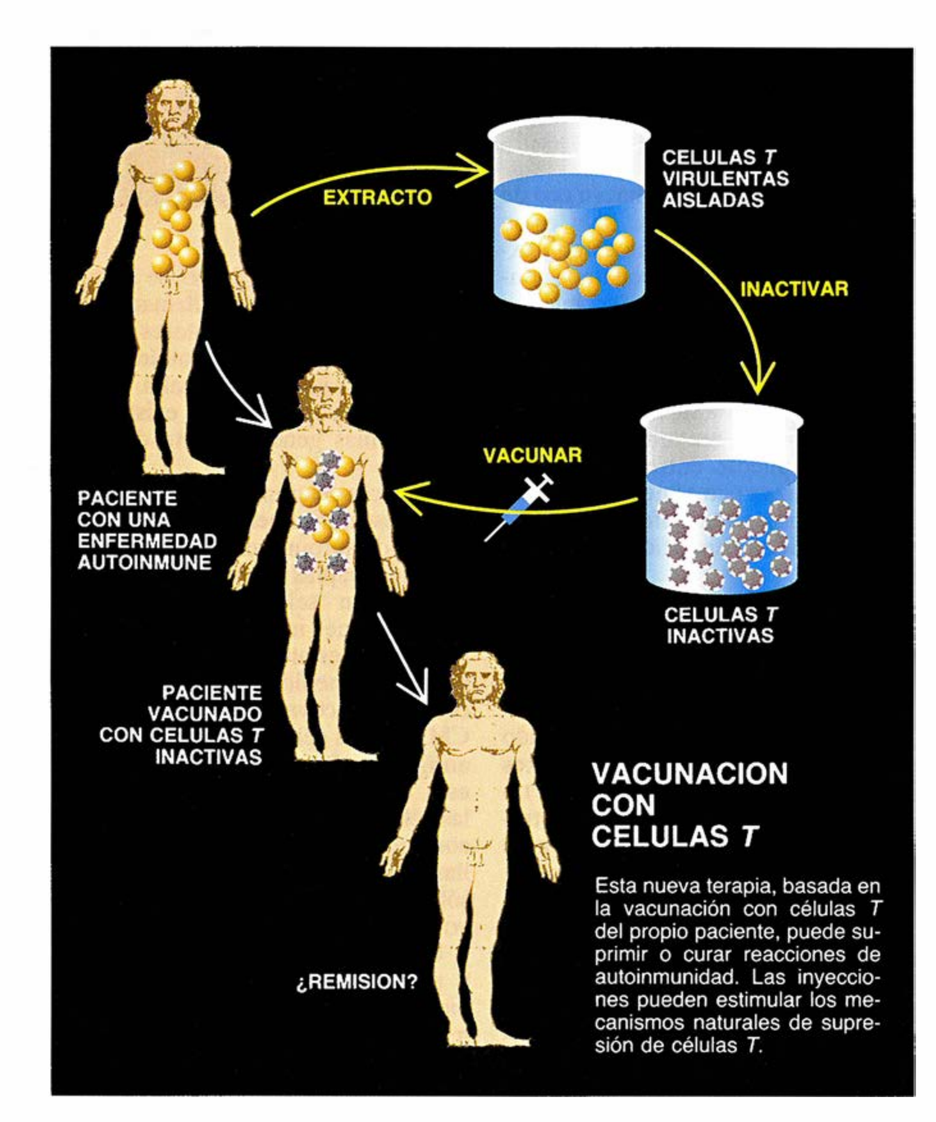
artificial que sirve de modelo de la EM en animales. Para inducir la EAE, los investigadores inyectan proteínas de mielina en ratones o ratas. El sistema inmune de los roedores, confundiendo la proteína con un antígeno, reacciona contra su propio sistema nervioso, con fatales consecuencias.

En 1988, dos grupos, uno dirigido por Steinman y el otro por Leroy E. Hood, demostraron, por separado, que las cadenas beta de los receptores antigénicos de las células *T* implicadas en la EAE eran sólo de dos tipos. Poco después, Steinman inyectaba anticuerpos contra uno de esos receptores en ratones con EAE.

Los resultados fueron espectaculares. La mayoría de los ratones que habían quedado parcial o totalmente paralizados por la EAE recuperaban su movilidad. Además, los anticuerpos al parecer también impedían el desarrollo de la EAE en muchos ratones a los que después se les inyectaba mielina. En junio del año pasado, Hood hizo público que su grupo había ido más lejos, inyectando una mezcla de dos anticuerpos contra los receptores de las células *T*, en ratones con EAE. De cinco ratones con EAE avanzada, cuatro se recuperaron; el tratamiento combinado parecía también prevenir la EAE en otros 19 de un grupo de 20. El único ratón que contrajo la EAE presentaba síntomas atenuados.

El rotundo éxito alcanzado con esos anticuerpos permite abrigar esperanzas de su aplicación en el hombre, para la EM y otras enfermedades de autoinmunidad. Pero antes habrá que contestar algunas preguntas. "Uno de los animales contrajo la enfermedad", recuerda Dennis M. Zaller, colaborador de Hood, "y sigo buscando el motivo". Zaller ha descubierto ya que dos series raras de células *T* estrechamente relacionadas, con diferentes receptores antigénicos, parecen propagarse y atacar el sistema nervioso.

"Creo –continúa Zaller– que los animales normales tienen también algunas de esas células *T*, pero ¿cuántas?". Es posible que una vez que el sistema inmune se haya sensibilizado contra un antígeno, algunos clones raros de células proliferen hasta llenar el hueco dejado por las que han sido neutralizadas con el tratamiento.



VACUNACION CON CELULAS *T*

Esta nueva terapia, basada en la vacunación con células *T* del propio paciente, puede suprimir o curar reacciones de autoinmunidad. Las inyecciones pueden estimular los mecanismos naturales de supresión de células *T*.

Vacunación activa

Llamar al orden a un sistema inmune errático que burla los tratamientos con los que se procuraba corregirlo puede acabar siendo un problema. Hafler y otros investigadores creen que la forma más prometedora de lograr una curación estable es ayudar al sistema inmune a que se mantenga él mismo en guardia, intensificando sus redes supresoras, si las tiene, o proporcionándole algunas.

Por ahora, la vía más eficaz para conseguir ese propósito es la vacunación activa, trabajo que iniciaron Cohen y sus colegas, en Israel, a principios de los años ochenta. Cohen demostró que podía curar ratas afectadas de EAE inyectándoles clones inocuos de células *T* que reaccionaban frente a la mielina extraídas de sus propios cuerpos; los clones habían quedado antes inactivados con radiaciones o productos químicos. Presumiblemente, la vacunación aumentaba la sensibilidad del animal a las células *T* antimielina causantes de la EAE, y sus sistemas inmunes suprimían la peligrosa reacción. "El sistema inmune parece receptivo a la vacunación con células *T*", dice Cohen, apoyándose en esos resultados para reivindicar la existencia de mecanismos supresores en el cuerpo.

Hace dos años, Hafler y Weiner iniciaron un estudio clínico de alcance limitado cuyo objetivo era comprobar la validez de esas ideas en el hombre. El estudio se hizo con cuatro pacientes afectados de EM; se trataba de ver si las vacunaciones activas eran lo suficientemente seguras. (Los pacientes se encontraban en los estadios más avanzados de debilidad provocados por la EM, y no respondían a los protocolos convencionales.) Tras hallar ciertos clones de células *T* que parecían haberse multiplicado en la sangre y sistema nervioso de los pacientes, cultivaron esas células en el laboratorio. Más tarde, las inactivaron con productos químicos y las volvieron a inyectar en los pacientes.

No se han observado efectos secundarios, aunque Hafler pone de relieve que el grupo estudiado es restringido en demasía para poder corroborar la calidad del tratamiento. A principios del año pasado, René de Vries, del Hospital Clínico de Leiden, comenzó un estudio similar con cuatro pacientes afectados de artritis reumatoidea. Utilizó clones de células *T* tomadas de sus articulaciones inflamadas.

A pesar de lo prometedor del tratamiento, la terapia basada en la vacunación activa con células *T* enteras no está exenta de inconvenientes. Producir la cantidad de células necesarias para el tratamiento es un proceso lento, caro y laborioso: hay que extraerlas de cada paciente, purificarlas, cultivarlas, inactivarlas y alterarlas químicamente para elevar su antigenicidad.

Los investigadores que trabajan en el campo de la vacunación activa están también interesados en depurar la técnica de usar péptidos (fragmentos de proteínas) específicos de células *T*, en vez de células completas. Estas moléculas podrían sintetizarse químicamente en grandes cantidades y administrarse a muchos pacientes.

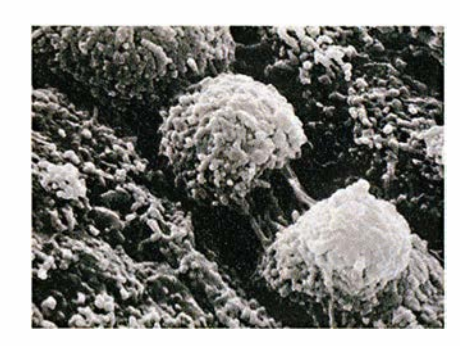
Steven W. Brostoff encabeza un grupo que investiga en la línea de la vacuna peptídica en la Inmune Response Corporation, de San Diego. Brostoff y sus colaboradores han fabricado pequeños fragmentos de péptidos de los receptores antigénicos encontrados en las células *T* que reaccionan frente a la mielina y causan la EAE. En un artículo publicado en noviembre de 1989, aseguraban que las ratas inyectadas con esos péptidos resistían la EAE.

Según Brostoff, tras la inyección las ratas desarrollan un cuadro de células *T* asesinas capaces de reconocer y suprimir selectivamente las células anti-mielina. Aunque la comercialización del producto está lejana, la compañía de Brostoff abordará este año los ensayos de toxicidad de los péptidos contra la artritis reumatoidea.

La idea de tratar enfermedades humanas de autoinmunidad con péptidos debe tomarse con un grano de sal. El verano del año pasado se analizaron datos todavía inéditos, que ponían de manifiesto sus riesgos potenciales. Según algunos investigadores, cuando inyectaban péptidos de células *T* en animales, los síntomas empeoraban. Nadie sabe qué es lo que hace al procedimiento bueno en unos casos y peligroso en otros. Mientras no se aclaren estas cuestiones, las aplicaciones clínicas tendrán que avanzar con suma cautela.

Hay otros dos enfoques sobre el tratamiento de la autoinmunidad. Se apoyan en los nuevos conocimientos adquiridos sobre los sistemas naturales de inducción de tolerancia. Ali Naji, de la Universidad de Pensilvania, exponía el suyo en septiembre del año pasado: inducir tolerancia a la insulina en ratas mediante una nueva técnica de trasplantes.

El grupo de Pensilvania trabajaba en un proyecto cuyo objetivo era superar el problema de los rechazos de órganos, esto es, la negativa del cuerpo a tolerar tejidos con proteínas del CMH foráneas. Se han sugerido los trasplantes como solución para los pacientes con diabetes insulinodependiente, cuyas células pancreáticas secretoras de la hormona han quedado destruidas por autoinmunidad. Desgraciadamente, aunque desde 1970 se intenta realizar trasplantes en seres humanos de células secretoras de insulina, los injertos celulares han sido siempre rechazados por sus receptores.





2. LA ARTRITIS REUMATOIDEA se produce cuando ciertos linfocitos *T* se infiltran en las membranas del fluido sinovial que recubren las articulaciones y las atacan (*izquierda*). Una micrografía de la cabeza del fémur de una cadera afectada (*derecha*), lisa de suyo, muestra profundas cavidades.

La heterodoxa solución de Naji a este problema consistió en trasplantar directamente células secretoras de insulina en el timo de animales de laboratorio, con la esperanza de "enseñar" al receptor a aceptar el injerto como algo propio. Después de insertar las células en el timo, administró una dosis de anticuerpos anti-células *T*, para eliminar las reservas de linfocitos y estimular una nueva ronda de producción de células *T*.

Los injertos de células secretoras en el timo perduraban indefinidamente. En apariencia, resistían los ataques inmunológicos. Además, la inserción de células secretoras similares en otras partes del cuerpo de algunas de esas ratas también era tolerada.

Naji cree que la presencia de las células foráneas en el timo estimula la deleción o inactivación de las células *T* que reconocen a los antígenos de los injertos y a los autoantígenos. El y sus colaboradores continúan sus estudios para comprobar si tales trasplantes pueden conferir tolerancia a tejidos rechazados por la autoinmunidad. Los resultados preliminares son alentadores.

Mientras tanto, Weiner y Hafler investigan un sistema muy distinto para vencer la autoinmunidad. Se basa en un fenómeno de tolerancia dirigida por el antígeno. Hace décadas que los inmunólogos se percataron de que la forma de exposición a un antígeno afectaba a la reacción del sistema inmune: los animales a los que se inyecta antígenos se tornan sensibles y alérgicos a los mismos; por el contrario, los animales que los toman con la dieta, los suelen tolerar. No somos alérgicos a las proteínas que comemos; sería muy perjudicial. Aunque nadie entiende exactamente cómo se origina la tolerancia en este caso, "utilizamos ese mecanismo natural para intentar suprimir las enfermedades de autoinmunidad", comenta Weiner.

En los últimos años, Weiner y sus colegas han demostrado que la administración de proteínas de mielina en la dieta de las ratas reduce la gravedad de la EAE inducida. Han tenido el mismo éxito suprimiendo ciertas formas de artritis y uveítis, una inflamación del ojo, inducidas experimentalmente. Weiner también ha llegado a inducir la supresión específica de antígeno, transfiriendo células *T* de animales tolerantes.

Los esfuerzos que se realizan para atajar las enfermedades de autoinmunidad avanzan, a pesar de la incertidumbre de las teorías que los orientan, y de la división entre las diferentes escuelas de inmunólogos. Resulta irónico que la rama de la inmunología que más tiene que ver con la tolerancia se haya vuelto contra una parte de sí misma, tratando a la teoría sobre la supresión con una actitud cercana al rechazo. La historia de este agitado campo demuestra que los conceptos que un año están pasados de moda, pueden convertirse en dogma al año siguiente.

BIBLIOGRAFIA COMPLEMENTARIA

LO PROPIO, LO AJENO Y LA AUTOINMUNIDAD. Irun R. Cohen en *Investigación y Ciencia*, págs. 22-31; junio de 1988.

INMUNOLOGICAL TOLERANCE: COLLABORATION BETWEEN ANTIGEN AND LYMPHOKINES. Gustav J. V. Nossal en *Science*, vol. 245, n.° 4914, págs. 147-153; 14 de julio de 1989.

FRONTIERS IN BIOTECHNOLOGY: TOLERANCE IN THE IMMUNE SYSTEM. *Science*, vol. 248, n.° 4961, págs. 1335-1393; 15 de junio de 1990.

Juegos matemáticos

Teorema navideño de Pierre de Fermat, explicado en un cuentecito endiablado

Ian Stewart

Era Nochematibuena. En la gélida oficina, atestada y diminuta, un reloj de pared movió inexorable sus manecillas hacia la vertical y su carillón hizo sonar un número perfecto de campanadas. Pachi Pifias secó el libro mayor con papel Möbius, que tenía una sola cara para que la tinta no traspasara al otro lado. Cerró el libro de contabilidad y lo repuso en su estante. Mañana iba a ser su día de vacación anual. ¡No tendría que trabajar durante las próximas factorial de 4 horas! Se puso el abrigo y se enrolló al cuello la bufanda, tan raídos ambos que su dimensión fractal era menor que 2. Pasó al salir junto a su patrón. "Feliz Matenavidad, señor Paniagua", dijo alegremente.

"¡Bah, paparruchas!", rezongó malhumorado el viejo. "Todo el comercio cierra el día de Matenavidad, Pachi. ¿Se da cuenta de lo que eso significa?"

"Pues significa que tenemos todo un día libre, señor Paniagua."

"Lo que significa, Pachi, es que mañana no tendremos clientes. Que será un día en que la caja de 'La Almoneda de Antiguallas Matemáticas' no va a ingresar ni un miserable céntimo."

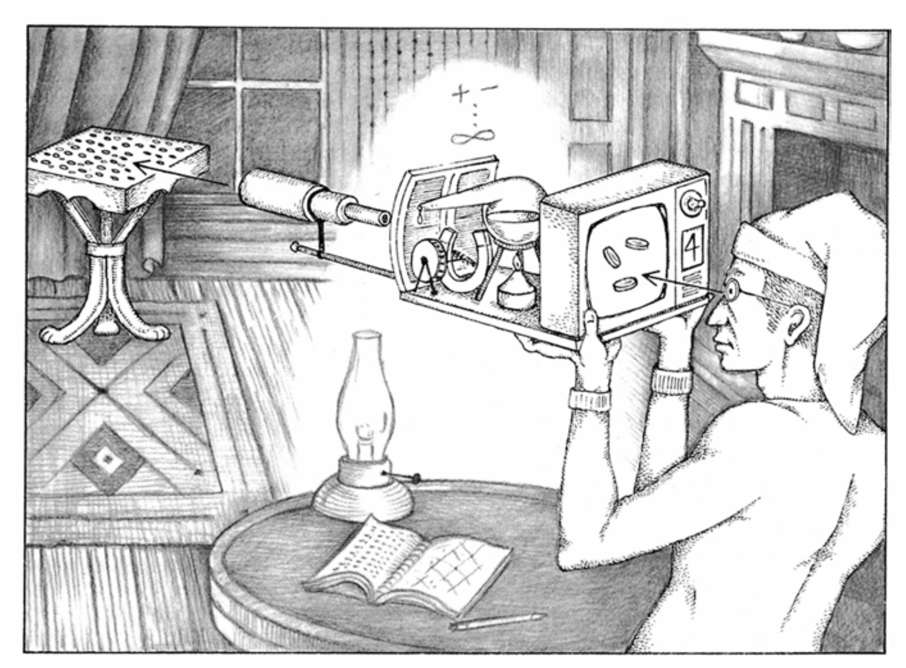
No podía decirse que el momento fuese el más oportuno, pero su esposa le había hecho prometer a Pifias que pediría un aguinaldo. "Esto... verá, Jefe..."

"Bueno, ¿qué pasa ahora?"

"Es que... me prometió un aguinaldo por Matenavidades, ¿se acuerda? Es para Gimotín, mi pequeño. El chiquillo no para de quejarse. Bastaría..."

"¿Un aguinaldo? ¡Pues no te digo! ¡Un aguinaldo! ¡Joven, una palabra más y le despido!"

Pifias se marchó, hundido y desmoralizado. Mas, para cuando llegó a su casa, el espíritu matenavideño había anulado su desesperación y le había infundido una visión más positiva de la vida.



1. *"Paniagua examinó a través del moduloscopio las 147 monedas de su mesilla de noche."*

"¿Quieres decir que no hay regalos?", chilló estridente el pequeño Gimotín.

"Tendremos que arreglarnos sin ellos, hijo mío", dijo Pifias, sintiendo que su animosa disposición decaía exponencialmente.

"¡Quiero un regalo! ¡Quiero un teorema nuevo! ¡O al menos, un lema de segunda mano! ¡Carlitos Mechero tiene uno que es súper! Por Dios, papá, hasta una mera conjetura sería mejor que nada!"

"¡No sabes cuánto lo siento, Gimotín, pero el señor Paniagua se guarda para sí sus conjeturas! Me temo que ni siquiera podemos disponer de una falacia pura y simple. Estoy en bancarrota total."

"Lo malo de ti, papá, es que no tienes ambición. Deberías pedir el puesto ese que ofrecen en Empeños Pitágoras!"

"Gimotín, no soy hombre altanero, pero jamás caeré tan bajo como para ir a vender triángulos rectángulos desgastados y con la hipotenusa vuelta del revés!" Pifias se impuso el deber de conservar la calma. "Mira, tendremos el pudín borracho que sobró de Pascua y lo vamos a pasar igual de bien que las Matenavidades de otros años. Y si tienes mucha, mucha suerte, a lo mejor logro desenterrar una vieja paradoja que me regaló tu madre. ¡Una buena mano de lógica y quedará como nueva!"

"¿Lógica intuicionista?", quiso saber el chiquillo, esperanzado. "¿No será una de las de siempre, trufada de verdadero y falso a palo seco?"

"¡Hijo mío, has tenido una idea luminosa!", convino Pifias. Gimotín se fue algo más apaciguado, y su padre se puso a buscar frenéticamente un enunciado cuya veracidad o falsedad no pudiera establecerse. Decidió telefonear al señor Paniagua para soli-

citarle un valor de verdad indeterminado, pero la telefonista le dijo que habían cortado la línea por falta de uso.

En un cochambroso apartamento de la otra punta de la ciudad, Bonifácil Paniagua procedía entre tanto a arrebujarse en su piltra, mientras danzaban por su mente los fantasmas de gastos e impuestos.

Se despertó cuando un viento helado enredó las cortinas y provocó el golpeteo de las hojas de la ventana. Saltó del lecho para cerrarla, pero se encontró con que ya lo estaba. ¿De dónde podría venir el viento...?

"Bonifáááácil", le llegó tétrica una voz. Paniagua se sumergió nuevamente en la cama y se hizo un ovillo bajo los cobertores. "¿Quién...quién es...?"

"Yo", dijo una voz tonante. "Soy el Espectro de los Teoremas Pasados. ¡He venido a llevarte conmigo, Paniagua!" El espectro tendió su etérea mano, que Paniagua asió de mala gana.

Paniagua se encontró súbitamente en una sala forrada en madera. Un hombre vestido con un ropón negro estaba escribiendo con una pluma de ave. "¿Dónde estamos?", preguntó Paniagua.

"En Francia. Es el día de Navidad de hace exactamente 350 años."

"¿Quién es el pelucón ese de ahí?"

"El caballero de la peluca, Paniagua, es el gran matemático francés Pierre de Fermat, que es recordado sobre todo por su célebre e indemostrado 'teorema magno', o 'último teorema'. Fue uno de los fundadores de la teoría de números. Está escribiendo una carta a su amigo Marin Mersenne. Si viajásemos de retorno al presente podríamos leer la carta original, que lleva fecha del 25 de diciembre de 1640. La carta da cuenta a Mersenne de un maravilloso descubrimiento."

"¿Y en qué consiste tan gran revelación?"

"Es conocida por teorema navideño de Fermat. Ciertos números primos son expresables como suma de dos cuadrados perfectos. Por ejemplo,

$$5 = 1 + 4 = 1^2 + 2^2$$

o

$$13 = 4 + 9 = 2^2 + 3^2.$$

Otros números primos, en cambio, no admiten tal descomposición, como por ejemplo el 3 o el 11. Pierre de Fermat descubrió cuáles la admiten y cuáles no."

Paniagua extrajo una sobada libreta y se puso a calcular. Al poco tenía resuelto el problema para todos los números primos menores que 100 [*véase la figura 2*].

"¿Se percata de la regularidad?", preguntó el Espectro de los Teoremas Pasados. Paniagua negó con la cabeza. "En vista de que es Matenavidad, le voy a dar dos pistas. La primera es que hay que despreciar el número 2, que es excepcional [siendo el único primo par, no es muy de extrañar que se salga de la norma]. La segunda consiste en fijarse en los restos que se obtienen al dividirlos por 4. Todo número impar es 1 unidad mayor que un múltiplo de 4, o es 3 unidades mayor; es decir, es de la forma $4k + 1$ o de la forma $4k + 3$. Por ejemplo, $5 = 4 \times 1 + 1$ es del tipo $4k + 1$."

Paniagua añadió a su tabla una nueva columna para poner de manifiesto si cada primo dado era de la forma $4k + 1$ o $4k + 3$, y al hacerlo salió a la luz la ley o regularidad. "Los números primos expresables como suma de dos cuadrados parecen ser de la forma $4k + 1$", dijo Paniagua, sorprendido. "Si se deja aparte el 2, que, según me dice, es excepcional."

"Excelente. Ahora bien, Fermat no se limitó a enunciar la conjetura, sino que la demostró. O al menos, esbozó un método para demostrarla." Y mientras el Espectro de los Teoremas Pasados se esfumaba y desaparecía, Paniagua pudo oírle murmurar, "Leonhard Euler dio una demostración definitiva allá por 1754..."

Paniagua se encontró de retorno en su helado dormitorio. Se esforzó por conciliar el sueño, pero el extraño teorema de Fermat no cesaba de martillear en su mente. Primos. Sumas de cuadrados. Restos de la división por 4. ¡La locura! Se revolvió en la cama, dio vueltas y más vueltas, entró a saco en el refrigerador. Pero no pudo dormir.

También Pachi Pifias se revolvía y agitaba en un lecho mucho menos cómodo, devanándose los sesos para imaginar cómo podría dar a tiempo con un valor de verdad indeterminado que depositar en el calcetín matenavideño de Gimo.

Paniagua llevaba dormido menos de $29 = 4 + 25$ segundos cuando oyó un pavoroso aullido y choque atronador. ¿Habría vuelto el vecino a lanzarle su zapatilla al gato de la viuda? No. El ruido se había producido en la propia habitación de Bonifácil. Empezó a temblar de miedo. Ante sus ojos se estaba materializando una figura luminosa.

NUMERO PRIMO	¿ES SUMA DE CUADRADOS?	¿ES 4k + 1 O 4k + 3?
2	$1^2 + 1^2$	excepción
3	no	(4 x 0) + 3
5	$1^2 + 2^2$	(4 x 1) + 1
7	no	(4 x 1) + 3
11	no	(4 x 2) + 3
13	$2^2 + 3^2$	(4 x 3) + 1
17	$1^2 + 4^2$	(4 x 4) + 1
19	no	(4 x 4) + 3
23	no	(4 x 5) + 3
29	$2^2 + 5^2$	(4 x 7) + 1
31	no	(4 x 7) + 3
37	$1^2 + 6^2$	(4 x 9) + 1
41	$4^2 + 5^2$	(4 x 10) + 1
43	no	(4 x 10) + 3
47	no	(4 x 11) + 3
53	$2^2 + 7^2$	(4 x 13) + 1
59	no	(4 x 14) + 3
61	$5^2 + 6^2$	(4 x 15) + 1
67	no	(4 x 16) + 3
71	no	(4 x 17) + 3
73	$3^2 + 8^2$	(4 x 18) + 1
79	no	(4 x 19) + 3
83	no	(4 x 20) + 3
89	$5^2 + 8^2$	(4 x 22) + 1
97	$4^2 + 9^2$	(4 x 24) + 1

2. *¿Qué números primos son suma de dos cuadrados?*

"Yo", dijo la figura con voz de trueno, "soy el Espectro de las Nociones por Venir".

"Espíritu detestable, obra a tu placer, pues estoy demasiado cansado para oponerme."

El fantasma de las nociones por venir depositó una caja sobre la mesa y ordenó: "¡Abrela!"

Paniagua vio en su interior lo que parecía ser una pantalla de televisión con un dial. La sostuvo en las manos. "¿Qué es?"

"Es un moduloscopio. Le permite a uno hacer caso omiso de aquello que en realidad no quiere ver."

"¿Como los pobres, por ejemplo? Bueno, de todos modos yo nunca les hago caso."

"¡Ah, pero con un moduloscopio uno deja de ver cosas de forma perfectamente sistemática. En particular, al sintonizar el dial en un número, si miramos por el dispositivo todos los múltiplos de ese número resultarán eliminados. Ponlo en el 'canal 4' y mira mis dos manos. ¿Cuántos dedos ves?"

"Dos. ¿Adónde han ido a parar los otros ocho?"

"El moduloscopio empieza restando primero ocho, porque 8 es el máximo de los múltiplos de 4 que es menor que 10; a continuación exhibe los

dos dedos restantes. Los matemáticos describen más sucintamente el funcionamiento del instrumento diciendo que '10 módulo 4 es igual a 2'. Paniagua examinó a través del moduloscopio las 147 monedas que había sobre su mesilla de noche. Al ver sólo tres soltó un chillido de rata. Apartó con presteza el moduloscopio de sus ojos y verificó, con gran alivio, que todas las monedas seguían en su sitio.

"Basta ya de tonterías", le espetó el fantasma. "Deja el dial en el canal 4, saca ese cuaderno tuyo y consulta la tabla de números primos. ¿Qué es lo que ves?"

"Nada excepto unos y treses... salvo el 2, que es excepcional. Cada número primo suma de dos cuadrados se ha convertido en un 1; todos los primos que no son suma de dos cuadrados han pasado a ser un 3. Aunque, claro, esta es la condición de ser $4k + 1$ o $4k + 3$; al tomarlos módulo 4 habrán de dar 1 o 3."

Quedó callado un momento. "De todos modos, sigo sin comprender qué interés tiene tomar el valor del primo, módulo 4."

"En lugar de observar los números primos, fíjate en los cuadrados."

Paniagua ojeó la tabla a través del moduloscopio. Se produjo un largo silencio. "Lo único que veo es la ecuación $1 = 0 + 1$, repetida una y otra vez."

"En efecto. ¿Y te das cuenta de por qué?"

"¿Será porque, módulo 4, los únicos cuadrados son 0 y 1?"

"Exactamente. Si elevamos al cuadrado un número par se obtiene un múltiplo de 4, que visto al moduloscopio se nos presenta como un 0. Al elevar al cuadrado los números impares va resultando 1, 9, 25, 49,... todos ellos una unidad mayores que un múltiplo de 4. Así pues, al tomar módulo 4 las sumas de cuadrados resulta $0 + 0 = 0$; $0 + 1 = 1$ o $1 + 1 = 2$. ¿Qué es lo que falta?"

"No veo el 3", dijo Paniagua.

"Correcto. Al tomar módulo 4 una suma de dos cuadrados, puede obtenerse 0, 1 o 2, pero nunca 3. Por consiguiente, los números primos de la forma $4k + 3$ (y en general, ningún número de esta forma) no pueden, de ningún modo, ser suma de dos cuadrados. Te das cuenta ahora de la importancia del módulo 4, ¿no es así?" El fantasma empezó a esfumarse lentamente.

"¡No te vayas!", gritó Paniagua en su desesperación. "Todo eso está muy bien, pero no demuestra que todos los primos de la forma $4k + 1$ sean suma de dos cuadrados, ¿verdad? Lo único que se ha probado es que los otros definitivamente no lo son."

La respuesta le llegó muy débilmente. "Tienes razón. Pero la salvación está al alcance de tu mano. Guarda bien el moduloscopio, y espeeeraaa..."

"Condenación", pensó Paniagua, "las desgracias siempre vienen de tres en tres. Tiene que ser otro espectro". Alzó la voz hacia el techo del cuarto. "¡Vamos, materialízate ya! ¡No tenemos toda la noche!"

"¡Yo soy el Fantas..atchíís!"

"¿Cómo dices?"

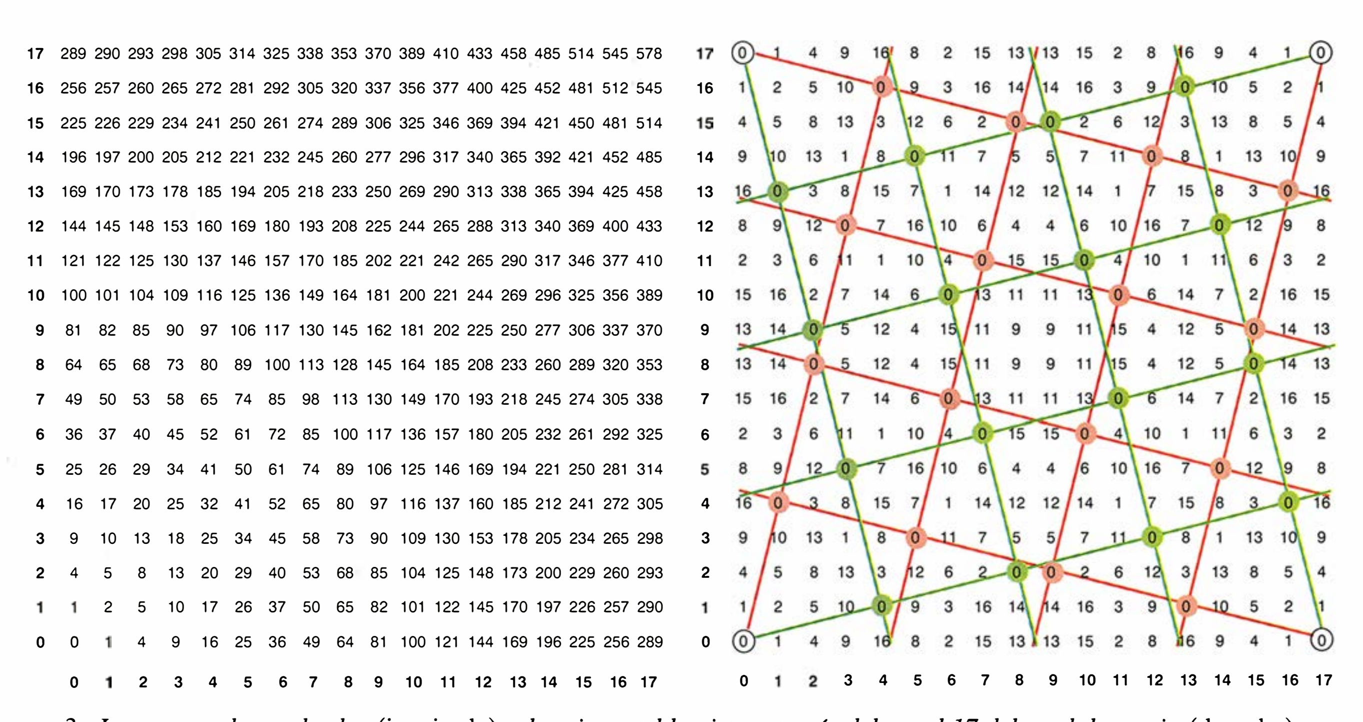
"El Fantasma de las Demostraciones del Presente. Oye, aquí hace un frío que pela. ¿Qué pasa, tronco, es que no enciendes nunca la chimenea?" El Fantasma se sonó ruidosamente la nariz en un jirón de ectoplasma.

"¿Has venido a decirme cómo demostrar que todo número primo de la forma $4k + 1$ es suma de dos cuadrados?"

"¡Diana, tío! Los fantasmas tenemos a veces extrañas ocupaciones. Pero no vamos a perder el tiempo. Sintoniza tu dial en el canal 17 y todo irá de perlas." Con una floritura, el Fantasma de las Demostraciones del Presente sacó una hoja de acetato dividida en casillas, y la colocó sobre la mesa. "Te lo voy a mostrar con el primo 17, pero el mismo método funciona en general. La idea, amigo mío, consiste en dejar de pensar en los números primos, y partir de las sumas de cuadrados. Esta hoja especial contiene todas las posibles sumas de dos cuadrados, ocupando $x^2 + y^2$ el puesto de la columna x y de la fila y. Obsérvala a través de tu moduloscopio. ¿Qué ves?"

"Un montón de números que van de 0 a 16, que ocupan toda la hoja."

"¡Vaya, se me pasó! Toma este rotulador y ve rodeando con un círculo cada 0 que veas, ¿vale?" Paniagua creó con los círculos un



17	289	290	293	298	305	314	325	338	353	370	389	410	433	458	485	514	545	578
16	256	257	260	265	272	281	292	305	320	337	356	377	400	425	452	481	512	545
15	225	226	229	234	241	250	261	274	289	306	325	346	369	394	421	450	481	514
14	196	197	200	205	212	221	232	245	260	277	296	317	340	365	392	421	452	485
13	169	170	173	178	185	194	205	218	233	250	269	290	313	338	365	394	425	458
12	144	145	148	153	160	169	180	193	208	225	244	265	288	313	340	369	400	433
11	121	122	125	130	137	146	157	170	185	202	221	242	265	290	317	346	377	410
10	100	101	104	109	116	125	136	149	164	181	200	221	244	269	296	325	356	389
9	81	82	85	90	97	106	117	130	145	162	181	202	225	250	277	306	337	370
8	64	65	68	73	80	89	100	113	128	145	164	185	208	233	260	289	320	353
7	49	50	53	58	65	74	85	98	113	130	149	170	193	218	245	274	305	338
6	36	37	40	45	52	61	72	85	100	117	136	157	180	205	232	261	292	325
5	25	26	29	34	41	50	61	74	89	106	125	146	169	194	221	250	281	314
4	16	17	20	25	32	41	52	65	80	97	116	137	160	185	212	241	272	305
3	9	10	13	18	25	34	45	58	73	90	109	130	153	178	205	234	265	298
2	4	5	8	13	20	29	40	53	68	85	104	125	148	173	200	229	260	293
1	1	2	5	10	17	26	37	50	65	82	101	122	145	170	197	226	257	290
0	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196	225	256	289
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

17	0	1	4	9	16	8	2	15	13	13	15	2	8	16	9	4	1	0
16	1	2	5	10	0	9	3	16	14	14	16	3	9	0	10	5	2	1
15	4	5	8	13	3	12	6	2	0	0	2	6	12	3	13	8	5	4
14	9	10	13	1	8	0	11	7	5	5	7	11	0	8	1	13	10	9
13	16	0	3	8	15	7	1	14	12	12	14	1	7	15	8	3	0	16
12	8	9	12	0	7	16	10	6	4	4	6	10	16	7	0	12	9	8
11	2	3	6	11	1	10	4	0	15	15	0	4	10	1	11	6	3	2
10	15	16	2	7	14	6	0	13	11	11	13	0	6	14	7	2	16	15
9	13	14	0	5	12	4	15	11	9	9	11	15	4	12	5	0	14	13
8	13	14	0	5	12	4	15	11	9	9	11	15	4	12	5	0	14	13
7	15	16	2	7	14	6	0	13	11	11	13	0	6	14	7	2	16	15
6	2	3	6	11	1	10	4	0	15	15	0	4	10	1	11	6	3	2
5	8	9	12	0	7	16	10	6	4	4	6	10	16	7	0	12	9	8
4	16	0	3	8	15	7	1	14	12	12	14	1	7	15	8	3	0	16
3	9	10	13	1	8	0	11	7	5	5	7	11	0	8	1	13	10	9
2	4	5	8	13	3	12	6	2	0	0	2	6	12	3	13	8	5	4
1	1	2	5	10	0	9	3	16	14	14	16	3	9	0	10	5	2	1
0	0	1	4	9	16	8	2	15	13	13	15	2	8	16	9	4	1	0
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

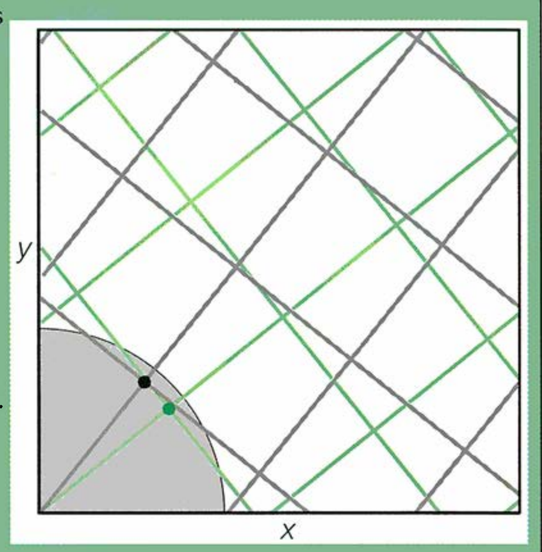
3. *Las sumas de cuadrados* (izquierda) *y la misma tabla vista a través del canal 17 del moduloscopio* (derecha).

Demostración de Minkowski del teorema de los dos cuadrados

Demuéstrese que si p es un número primo de la forma $4k + 1$, entonces p se representa como una suma de dos cuadrados.

Supongamos que p sea un número primo de la forma $4k + 1$.

Hallemos todos los puntos (x,y) tales que $x^2 + y^2$ sea múltiplo de p.



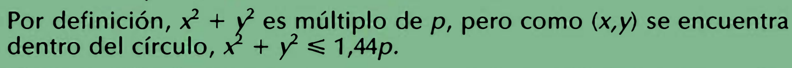
Elijamos uno de los dos retículos formados por puntos (x,y).

El área de los paralelogramos de cada uno de los retículos es p unidades cuadradas.

Tracemos un círculo, con centro en el origen, de radio suficientemente mayor que $\sqrt{p}$; por ejemplo, radio $= 1{,}2\sqrt{p}$.

El área del círculo es $1{,}44\pi p \approx 4{,}52p$, que es mayor que $4p$.

Por el teorema de Minkowski, el círculo contiene un punto del retículo distinto del (0, 0).



Por definición, $x^2 + y^2$ es múltiplo de p, pero como (x,y) se encuentra dentro del círculo, $x^2 + y^2 \leqslant 1{,}44p$.

El único múltiplo no nulo de p que es menor o igual que $1{,}44p$ es el propio p, porque el múltiplo siguiente es $2p$, demasiado grande.

Así pues, $x^2 + y^2 = p$ exactamente, y el teorema está demostrado.

curioso motivo periódico [*véase la figura 3*]. Lo observó sin gran entusiasmo, y acabó meneando negativamente la cabeza.

"Hay una pauta oculta", dijo el espectro. "Voy a colorear unos círculos de rojo, y otros de verde. ¿Observas algo?"

"¡Quién lo diría! Son sencillamente dos cuadriculados regulares, uno encima de otro."

"¡Correcto! El nombre adecuado para cuadriculados como éstos es 'retículo'. Los puntos coloreados que hemos señalado constituyen la totalidad de los puntos (x, y) tales que $x^2 + y^2$ sea múltiplo de 17. Ahora, fíjate en el retículo rojo y dime cuál es el punto más cercano al origen (columna 0, fila 0)."

"Eso es fácil. Se encuentra en la columna 1, fila 4."

"Y la correspondiente suma de cuadrados, ¿qué múltiplo de 17 es?"

"¡$1^2 + 4^2$ es igual al propio 17! ¡Ya caigo! Me estás diciendo que cada punto del retículo rojo que se encuentre a mínima distancia del origen resuelve el problema de representar 17 como suma de dos cuadrados!"

"Y mucha razón que tienes. El retículo verde también funciona, pero sólo nos da la solución $4^2 + 1^2$, que es la misma en orden inverso. Prueba con otro número: sintoniza esta vez el aparato en el canal 41. ¿Vuelve a ocurrir lo mismo?"

"¡Sí, mira, vuelven a aparecer dos retículos superpuestos!", dijo Paniagua. "Y el punto más próximo al origen de la retícula roja está en la columna 4 y la fila 5, y $4^2 + 5^2 = 41$!"

"Fenomenal. Puedes divertirte mucho probando con otros canales de tu moduloscopio. Elige un primo p y señala todas las posiciones (x, y) en las cuales $x^2 + y^2$ sea múltiplo de p. Se obtiene siempre una configuración compuesta por dos retículos, aunque probablemente no te habrías dado cuenta si no te hubieran dicho que la buscaras.

"¡Pero yo soy el Fantasma de las Demostraciones Presentes, y no sólo de los Ejemplos! Tengo que explicar por qué hay dos retículos y por qué el punto del retículo más cercano al origen siempre resuelve el problema. Veamos primero la existencia de dos retículos. Eso se debe a las raíces cuadradas de -1."

"No sabía que -1 tuviera raíces cuadradas", le cortó Paniagua.

"Ah, ya. Bueno, ningún número real puede tener cuadrado igual a -1, así que se inventó un nuevo número, llamado i, tal que $i^2 = -1$; así nació el sistema de los números complejos. Pero si se tiene moduloscopio, no hay necesidad de complejos." Anotó algo en la hoja de plástico. "Echa una ojeada a esto a través del moduloscopio, todavía en el canal 17." Paniagua miró.

La hoja decía:

$$x^2 + y^2 = (x + 4y)(x - 4y)$$

"Parece un disparate, ¿verdad? Pero casi todas las cosas resultan un poco absurdas vistas al moduloscopio... hasta que se las interpreta debidamente. Sabemos por álgebra elemental que $(x + 4y)(x - 4y) = x^2 - 16y^2$. Pero a través del moduloscopio -16 es lo mismo que $17 - 16$ (porque los múltiplos de 17 son invisibles), diferencia que es precisamente 1. Por tanto, en el moduloscopio, $x^2 - 16y^2 = x^2 + y^2$.

"Los puntos que señalaste", hizo notar el fantasma, "son aquellos que, vistos por el moduloscopio, satisfacen la ecuación $x^2 + y^2 = 0$, ecuación que módulo 17 se factoriza en $(x + 4y)(x - 4y) = 0$; o sea: $x = -4y$ o $x = 4y$. Cada factor corresponde a uno de los retículos. El retículo rojo está dado por la ecuación $x = -4y$ y el verde, por la $x = 4y$. ¡Siempre módulo 17, claro está! Observa todos los retículos y verifícalo. Por ejemplo, en el retículo verde se encontrará que todos los puntos, como (4,1), (8,2), (12,3), (16,4) y demás, verifican la ecuación $x = 4y$.

"Esa es la primera idea importante. Al observarse por el moduloscopio, sintonizado en el canal 17, el número -1 sí tiene raíz cuadrada, que es 4! En efecto, $4^2 + 1 = 17 = 0$. Y ello nos lleva directamente a la existencia de los dos retículos. Ocurre exactamente lo mismo para cualquier primo que sea de la forma $4k + 1$. Según resulta, esos primos son precisamente aquellos para los que -1 tiene raíz cuadrada. ¿Estás preparado para la segunda idea importante?"

"Tan dispuesto como jamás podré estarlo", replicó Paniagua.

"Cada retículo está compuesto por paralelogramos idénticos. [En este caso, los paralelogramos son en realidad cuadrados un poco inclinados, pero en otros retículos no ocurre así, por lo que seguiré llamándolos para-

lelogramos.] ¿Qué área tiene uno de estos paralelogramos? Haz la prueba en algunos ejemplos."

Bonifácil garabateó en su cuaderno y respondió:

"Cuando $p = 17$, el paralelogramo tiene entonces un área de 17 unidades cuadradas. Y cuando $p = 41$, el área es de 41 unidades cuadradas. Supongo que, para un primo cualquiera p, el paralelogramo tendrá un área de p unidades".

"Así es, en efecto, pero no me pararé a demostrarlo. Ahora, seguramente estarás preguntándote a qué se debe mi interés por el área del paralelogramo."

"Se me había pasado por la imaginación."

"Se debe a un teorema demostrado por Hermann Minkowski, matemático ruso que ejerció docencia en Alemania. Fue el inventor del espacio de Minkowski, utilizado por Einstein en su teoría de la relatividad. Minkowski tuvo una brillante idea sobre retículos, cuya aparente simplicidad es engañosa: si el área del paralelogramo es sumamente pequeña, los puntos del retículo habrán de encontrarse muy próximos. Así que algunos de ellos habrán de hallarse cercanos al origen.

"Minkowski demostró un teorema en el que daba forma precisa a esta idea. Supongamos que se tenga un retículo formado por paralelogramos y un círculo con centro en el origen. El teorema de Minkowski dice que, siempre que el área del círculo sea al menos cuatro veces mayor que la del paralelogramo, el círculo contendrá en su interior puntos de retículo distintos del origen.

"Podemos servirnos del teorema de Minkowski para demostrar que el punto reticular más cercano al origen resuelve la representación de p en suma de dos cuadrados. Sea el ejemplo $p = 17$. El área del paralelogramo también es 17. Tomemos un círculo de radio algo mayor que la raíz de 17; por ejemplo, de radio 5. Su área es entonces $5^2\pi = 25\pi = 78{,}54$, que es mayor que $4 \times 17 = 68$, por lo que el teorema de Minkowski es aplicable. ¿Me sigues hasta aquí?"

"Estoy pendiente de tus palabras."

"Según el teorema de Minkowski, el círculo contiene al menos un punto reticular distinto del origen. Sea (x,y) tal punto. Entonces $x^2 + y^2$ es menor o igual que 25. Para puntos del retículo, $x^2 + y^2$ es múltiplo de 17. Tal múltiplo es no nulo, pues el punto reticular no es el origen. ¿Qué múltiplos no nulos de 17 son menores o iguales que 25?"

"El propio 17 y ninguno más", dijo Paniagua.

"¡Ya lo has cazado! Así que $x^2 + y^2$ tiene que ser exactamente 17, lo que resuelve el problema. El mismo método funciona en general", se jactó el fantasma. "La idea de Minkowski hizo nacer una nueva rama de las matemáticas, conocida por geometría de los números a causa de su libro sobre el tema, publicado en 1896. Se vale de la geometría para estudiar teoría de números, dos materias que normalmente uno no creería relacionadas. Otra aplicación de la geometría de números es el teorema de los cuatro cuadrados: todo número entero positivo (sea primo o no) es suma de cuatro cuadrados perfectos. Pero dejaré que ese problema te mantenga perplejo hasta las próximas Matenavidades, Bonifácil."

Paniagua pudo finalmente relajarse, y el torbellino en que se habían convertido sus ideas fue apaciguándose. Al amodorrarse se acordó de Pachi Pifias. Hallándose de humor particularmente benigno, prometió ser en el futuro más amable con su empleado.

Gimotín se despertó muy animado la mañana de Matenavidad. "¡Papá, papá! ¿Me has podido traer la vieja paradoja de mamá pintada con una mano fresca de lógica intuicionista, verdad? ¿La que tiene indeterminado su valor de verdad?"

"Verás, dijo Pachi. "Resulta más bien difícil de explicar, hijo." Le pasó una caja de cartón muy baqueteada, esforzándose en pronunciar las palabras que lograsen acallar los taladrantes chillidos del pequeño Gimotín. "No estoy seguro de haberlo hecho o no."

Fue un hallazgo genial. O quizás, el Fantasma de las Intuiciones Venideras, posado en su hombro, se las había susurrado al oído. Porque el rostro de Gimotín se encendió como un árbol de Matenavidad. "¡Gracias, papá! ¡Me la trajiste!"

Afrontémoslo: mal podría hallarse un valor de verdad más indeterminado que la respuesta de Pachi Pifias.

BIBLIOGRAFIA COMPLEMENTARIA

THE THEORY OF ALGEBRAIC NUMBERS. Harry Pollard. Mathematical Association of America, 1950.

MATHEMATICAL THOUGHT FROM ANCIENT TO MODERN TIMES. Morris Kline. Oxford University Press, 1972.

ALGEBRAIC NUMBER THEORY. Ian N. Stewart y David O. Tall. Chapman and Hall, 1987.

Libros

Combustión, flora regional, matemática, demografía, oceanografía y Eriúgena

José Luis Castillo y F. Javier de la Rubia, Josep M. Camarasa,
Javier Echevarría, Carlos Gracia y Luis Alonso

DISSIPATIVE STRUCTURES IN TRANSPORT PROCESSES AND COMBUSTION. Dirigido por D. Meinköhn. Springer Verlag; Berlín, 1990.

Desde que en 1977 el profesor Hermann Haken (director de la colección) inaugurara con su libro *Synergetics. An Introduction* la prestigiosa Springer Series in Synergetics, se han publicado ya casi cincuenta volúmenes entre monografías y actas de congresos y reuniones científicas. El indudable éxito de la serie se debe, en gran medida, al carácter interdisciplinar de la sinergética, entendida como el estudio de los fenómenos producidos por la cooperación de las partes individuales de un sistema. Este carácter interdisciplinar conduce a que un buen número de los temas activos de investigación en la actualidad en disciplinas tan dispares, aparentemente, como la física, la química, la biología e incluso la sociología tengan cabida dentro del campo de la sinergética.

El libro que nos ocupa hace el número 48 de la serie y recoge parte de las conferencias presentadas en el seminario que, sobre el tema de la combustión, de importancia tanto práctica como teórica, se celebró en Bielefeld (Alemania) del 17 al 21 de julio de 1989. Antes que nada hay que decir que el hecho de que el número de los trabajos recogidos en el libro sea reducido, y, por consiguiente, la extensión de cada uno de ellos más amplia de lo habitual, hace que su utilidad sea mayor que la de los libros de actas al uso, pues no sólo se puede saber *qué* se hace sino, también, *cómo* se hace en el tema de investigación que nos ocupa.

El contenido del libro está dividido en cuatro partes. La parte I *(Fluid Flow: Turbulence, Modelling)* incluye dos capítulos sobre modelos de flujos no reactivos. El primero, de S. Grossmann, resume los últimos avances en el estudio de la turbulencia homogénea isótropa totalmente desarrollada, presentada como una solución caótica determinista de las ecuaciones de la dinámica de fluidos. Mediante una descomposición de la velocidad en un término de campo medio, a gran escala, y una subescala a la que contribuyen los remolinos más pequeños (a diferencia de la clásica descomposición de Reynolds en un valor medio y una parte fluctuante), la turbulencia se estudia a partir de la definición de una función de estructura de la que dependen la viscosidad turbulenta y las correlaciones espacio-temporales.

El segundo capítulo, de R. Nasilowski, trata de un autómata celular como modelo de gas discreto. Este tipo de autómatas celulares aparecen, en la actualidad, como una alternativa para la simulación de flujos complejos, frente a la resolución mediante métodos numéricos de las ecuaciones de movimiento fluido. Tienen como principal ventaja la sencillez de su realización y no requieren tanta capacidad de memoria ni de cálculo como los tradicionales métodos de diferencias finitas o elementos finitos. Este capítulo introduce un autómata que corre sobre una red formada por dos redes cuadradas intercaladas, alternándose las partículas del autómata entre una red y otra para cada paso temporal. Las reglas de interacción (colisiones entre partículas en un nodo) son totalmente deterministas. Se demuestra que, para una determinada densidad de partículas, las ecuaciones de evolución coinciden con las de un fluido de Euler (no viscoso).

La parte II *(Combustion Waves, Flames)* es bastante homogénea en su contenido. En su capítulo, G. Joulin deduce fenomenológicamente los términos dominantes en la ecuación de evolución no-lineal de un frente de llama moderadamente deformado, obteniendo la clásica ecuación de Michelson-Sivashinsky. Mediante una descomposición en polos demuestra que la ecuación admite soluciones periódicas y aperiódicas. Utiliza el método para el estudio de un frente inicialmente plano mostrando su inestabilidad y la forma estable a que tiende el frente deformado; la solución analítica coincide con los últimos resultados numéricos obtenidos. Estudia también el frente inicialmente parabólico mostrando la aparición de inestabilidades secundarias que son emitidas desde el vértice y se transmiten a lo largo de la parábola aumentando en el trayecto su longitud de onda. El capítulo incluye una muy completa y actualizada lista de referencias sobre el tema.

Por su lado, G.I. Sivashinsky en su capítulo estudia la velocidad efectiva de propagación de una llama en un medio turbulento con respecto a la velocidad de propagación laminar. Para ello simula la turbulencia como una cascada (superposición) de torbellinos isotrópicos y estadísticamente autosimilares con escalas espacio-temporales muy separadas. Los resultados muestran una buena concordancia con los experimentos para turbulencias cuya intensidad no es muy grande en relación a la velocidad de propagación laminar de la llama.

El siguiente trabajo, de N. Peters y Ch. Franke, tiene un gran valor didáctico. De forma concisa y con suma claridad se examina la interacción entre turbulencia y un frente de propagación de llamas. Los diferentes regímenes de combustión turbulenta de gases premezclados se caracterizan por las distintas escalas espacio-temporales que intervienen. La razón entre la velocidad de propagación turbulenta y la velocidad de un frente plano es igual a la relación entre la superficie del frente de llama turbulento y la del frente de llama plano. Se presenta un estudio de la dimensión fractal de la superficie de la llama y se compara con los resultados experimentales de la velocidad de propagación turbulenta. La concordancia entre el modelo y los resultados experimentales es prometedora. Un trabajo de M. Konczalla en el que se usa

el método de desarrollos asintóticos de Liñán, para llamas de difusión, en el estudio de la localización de una llama junto a una zona de recirculación, y otro de S.U. Schüffel en el que se estudian numéricamente las oscilaciones celulares transversas que se producen durante las detonaciones, cierran esta parte del libro.

La parte III *(Turbulent Combustion)* contiene dos únicos capítulos. El primero, de J. Warnatz, presenta un estudio de llamas de premezcla en flujos turbulentos con especial atención a los procesos químicos mediante la introducción de toda la complejidad de la reacción química. Hasta 72 reacciones elementales con 15 especies químicas se tienen en cuenta para el estudio de una llama de difusión CO-aire. Los perfiles de concentración de especies que se obtienen presentan gran concordancia con las medidas experimentales. Por su parte, en la contribución de H. Bockhorn, haciendo uso de un modelo de combustión turbulenta que incorpora toda la complejidad de la reacción química y, por medio de un análisis de sensitividad, se determinan qué reacciones elementales tienen mayor importancia en el cálculo de las fracciones másicas de hidrógeno en una llama turbulenta de difusión aire-hidrógeno.

La parte IV *(Reaction-Difussion Systems: Stability, Pattern Formation, Autowaves, Chaos)* es la más extensa del libro (ocupa más de la mitad del mismo) y contiene siete comunicaciones. En la primera de ellas, M. Bestehorn y H. Haken analizan, en un largo e instructivo trabajo, la formación de estructuras espacio-temporales en diversos sistemas muy alejados del equilibrio termodinámico. Partiendo de las ecuaciones básicas para cada sistema y mediante la eliminación adiabática de las variables no relevantes (es decir, eliminando las variables que relajan a sus valores estacionarios rápidamente y conservando las que evolucionan más lentamente) obtienen una ecuación generalizada de Ginzburg-Landau, con dependencia espacio-temporal, para el "parámetro de orden", o variable relevante, cerca del punto de bifurcación o de cambio de estabilidad, que, posteriormente, es analizada numéricamente para los distintos sistemas (reacciones químicas oscilantes, inestabilidad convectiva, mezclas binarias) mostrando cómo aparecen las distintas estructuras espacio-temporales observadas experimentalmente.

La aparición de oscilaciones y soluciones caóticas en reacciones químicas se considera en otros dos trabajos. P. Gray y S.R. Kay analizan, de forma teórica, los cambios (con relación al caso clásico de reacciones isotermas) en el comportamiento temporal que aparecen al considerar la variación de la temperatura en una reacción química (existencia de oscilaciones complejas y caos), mientras que F.W. Schneider y A.F. Münster presentan un detallado estudio experimental de la aparición de soluciones caóticas en la conocida reacción de Belousov-Zhabotinsky. Otros trabajos, de menor interés, sobre descripción cinética de procesos oscilatorios y movimiento hidrodinámico, uso de autómatas celulares isótropos para simulación de medios excitables, comparación de inestabilidades morfológicas y de Rayleigh-Marangoni y la aplicación de las ideas de la teoría de catástrofes a los fenómenos de ignición y extinción que aparecen en la combustión completan esta última parte.

En resumen, un interesante libro sobre algunos de los problemas actuales en el tema de la combustión, tratados con los métodos modernos utilizados para analizar el comportamiento sinérgico de los sistemas no lineales muy alejados del equilibrio termodinámico. (J. L. C. y F. J. R.)

FLORA MANUAL DELS PAÏSOS CATALANS, por Oriol de Bolòs, Josep Vigo, Ramon M. Masalles y Josep M. Ninot. Editorial Pòrtic S.A.; Barcelona, 1990.

Si es lícito hablar de géneros en la literatura científica, no hay duda de que el de las floras sería uno de los más firmemente establecidos. Esbozado en las *historiae plantarum* seiscentistas, normalizado por Linné y por sus discípulos, renovado por Lamarck y de Candolle, pujante a lo largo del siglo XIX y los primeros decenios del XX, ha experimentado en Europa un nuevo renacimiento, tras algunos decenios de eclipse, a partir de los años sesenta, gracias al esfuerzo prolongado e innovador de la publicación de *Flora Europaea* (1969-80), que contó con la colaboración de un considerable número de botánicos.

Una flora es una obra que trata del conjunto de plantas vasculares de un territorio definido, las enumera, las describe e indica cuál es, dentro de ese territorio, su hábitat preferente, su fenología, su abundancia o rareza, su distribución geográfica o altitudinal y otros datos de interés. Pero además, por lo menos desde la publicación por Lamarck de la primera edición de su *Flore Française* en 1778, una flora es, no sólo una enumeración, lo más exhaustiva posible, de las plantas vasculares de un territorio dado, ordenadas conforme a una sistemática definida y con informaciones pertinentes a cada una de ellas, sino también un conjunto de claves de determinación que permiten la identificación de cualquier ejemplar de una planta vascular que se ha hallado en tal territorio, asignándolo a una de las especies que figura en la flora (a menos, claro está, que se trate de un ejemplar de alguna especie que haya escapado hasta entonces a las pesquisas de los botánicos o que se haya introducido recientemente).

Si el primer aspecto, catálogo de una parte harto significativa del patrimonio natural, limita su interés a los profesionales de la botánica y de las ciencias afines y de la gestión del patrimonio natural, el segundo –claves de determinación– concierne a un conjunto mucho más amplio: docentes, estudiantes, naturalistas aficionados, excursionistas y muchos otros que no siempre se sienten cómodos ante una obra en varios gruesos volúmenes. Hasta ahora, eran inexistentes en España, si exceptuamos algunas guías de árboles y de otras plantas leñosas y flórulas locales de área muy restringida. Con la obra que comentamos, los autores vienen a satisfacer con creces esta demanda para el territorio de los "Países Catalanes" (Andorra, Baleares, Cataluña, la Franja oriental de Aragón catalanoparlante, Valencia y el departamento francés de los Pirineos Orientales) y, con algunas limitaciones, para un territorio que cubre prácticamente más de la mitad de la península Ibérica. Describe esquemáticamente las casi 3600 especies que enumera, facilita su identificación mediante las adecuadas claves dicotómicas y proporciona una información básica, aunque sucinta, de la significación ecológica y geográfica de cada especie, con indicación de su forma vital, de su fenología y, en los casos pertinentes, de su uso y aprovechamiento, de sus propiedades tóxicas o medicinales, etc.

Dado el carácter de la obra, la ilustración se limita a la representación de elementos o partes significativos de algunas plantas de identificación más difícil; aun así, el libro alcanza un grosor (más de 1200 páginas) y un peso (un kilo y medio) que lo hacen algo menos manejable de lo que sus autores –y más aún sus lectores– desearían, en particular como instrumento de trabajo de campo. Aunque concebida como enumeración, descripción y conjunto de claves de determinación de las especies que se dan en la zona, esta interesante *Flora manual dels Països Catalans*

está llamada a desempeñar el mismo papel para buena parte del resto de España, con exclusión quizá de las comarcas norteñas de carácter más atlántico y del extremo sur peninsular (y, por descontado, de las islas Canarias). Efectivamente, si comparamos su contenido con el de los dos primeros volúmenes (los únicos publicados) de la *Flora Ibérica,* veremos que sólo una veintena de los dos centenares de géneros que aparecen en ésta se hallan ausentes de la *Flora manual dels Països Catalans* y que, en casi todos los casos, se trata de géneros poco frecuentes, por lo común monoespecíficos (al menos en lo que respecta a su representación en la península Ibérica), de los que más de un tercio corresponden a especies introducidas y de los que todos, excepto cuatro, tienen su área restringida a las regiones peninsulares antes citadas.

Los autores han tenido un cuidado exquisito (ya tradicional entre nuestros botánicos desde Pius Font i Quer) en la elección del vocabulario, de modo que, al margen de pequeñas diferencias ortográficas, no resulta nada difícil reconocer las raíces griegas o latinas de cada vocablo técnico y con ellas su equivalente castellano y su significado. Alguna mayor dificultad, siempre de escasa entidad, puede ofrecer algún término de vocabulario común, tal adverbio o cual preposición, pero en ningún caso insuperable para el lector. (J. M. C.)

HISTORY AND PHILOSOPHY OF MODERN MATHEMATICS. Dirigido por William Aspray y Philip Kitcher. Minnesota Studies in the Philosophy of Science; University of Minnesota; Minneapolis, 1988.

La reflexión sobre la ciencia ha estado centrada durante el siglo XX en las teorías físicas (Newton, Einstein, mecánica cuántica, etc.), biológicas (Darwin, Mendel, Wilson...) y en general en las ciencias empíricas. Por influencia del positivismo lógico, del formalismo de Hilbert y del falsacionismo de Popper, las matemáticas han merecido hasta los años setenta muy escasa atención. La línea de investigación iniciada por Frege y continuada por Russell, von Neumann, Gentzen y muchos otros, que trataba de buscar una fundamentación lógica de las matemáticas, ha dominado a lo largo de muchos años, tendiendo a identificar la filosofía de las matemáticas con la metamatemática.

En los últimos años se advierte una clara inflexión en dicha tendencia. Sin olvidar la labor pionera de Polya, cuyas ideas resurgen con fuerza, cabe señalar a Kuhn y a Lakatos, junto con la publicación de los escritos del segundo Wittgenstein sobre filosofía de las matemáticas, como los motores de esta renovación del pensamiento en torno a las matemáticas. En los últimos años se publican con frecuencia obras que replantean la identificación kantiana de las matemáticas con un saber a priori y promueven nuevas corrientes de pensamiento, que van desde la reaparición del empirismo en matemáticas hasta la revalorización del constructivismo, pasando siempre por una estrecha conexión entre la filosofía de las matemáticas y la historia de las matemáticas.

El libro que comentamos es un buen ejemplo de dicha tendencia. Retomando la idea de Birkhoff al organizar un simposio en 1974 en la Universidad de Harvard sobre estas cuestiones, en el que se convocó conjuntamente a historiadores y filósofos de las matemáticas, la Universidad de Minnesota celebró una Conferencia en mayo de 1985 con presencia de destacados especialistas, resultado de la cual es la presente obra, que viene precedida por una utilísima introducción (pp. 3-57) a cargo de los editores, Aspray y Kitcher. Las actas del simposio de 1985 incluyen las contribuciones de autores tan importantes como Askey, Aspray, Bennett, Birkhoff, Browder, Crowe, Daston, Dauben, Edwards, Friedman, Goldfarb, Grabiner, Kitcher, Moore y Stein: una bien seleccionada representación, sin duda, de la actual vanguardia investigadora en estas materias.

La "opinionated introduction" de Kitcher y Aspray plantea admirablemente bien el estado de la cuestión. Tras compendiar el desarrollo de la filosofía de las matemáticas en el siglo XX, partiendo de Frege y de las tres opciones que se derivan de su programa de fundamentación lógica de las matemáticas (logicismo, formalismo e intuicionismo), los editores van indicando las sucesivas críticas a dicho programa, hasta llegar a la obra fundamental de Lakatos, *Pruebas y Refutaciones* (1963-64). A partir de este momento cabe señalar la existencia de dos programas de investigación en filosofía de las matemáticas: el primero, de estilo analítico, que investiga los fundamentos de la matemática, y el segundo, mucho más interesado en cuestiones históricas, que prescinde de la tarea de fundamentar las matemáticas, para preocuparse por la indagación de la metodología matemática. Tal y como Kitcher y Aspray indican, su panorámica está basada en el ámbito anglosajón, prestando escasa atención a los estudios publicados en otras lenguas.

Pese a esta insuficiencia, la introducción resulta muy clara y muy útil, tanto para resumir la evolución de la filosofía de las matemáticas en el siglo XX, como sobre todo para destacar la importancia de la renovación producida en los EEUU y en Gran Bretaña con la profesionalización de los historiadores de la matemática: Dauben, Grattan-Guinness y May son destacados especialmente al respecto (p. 26). Desde el punto de vista institucional, se subraya la aparición de sociedades dedicadas a estos estudios en los EEUU, en Canadá y en el Reino Unido, así como la publicación de revistas como *Historia Mathematica, Archive for History of Exact Science, History and Philosophy of Logic, Annals of the History of Computing, Mathematical Intelligencer* y el *Bollettino di storia della scienze mathematiche,* en torno a las cuales se produce esta renovación y cambio. La sistemática publicación de las obras completas de los grandes matemáticos en los últimos años habría sido, asimismo, un importante factor de cambio, al permitirnos conocer las vías concretas mediante las cuales llegaron a sus descubrimientos, en lugar de limitarnos a conocer sus teorías y sus contribuciones a través de las reconstrucciones que ofrecen los libros de texto, basadas siempre en una concepción acumulativa del progreso en matemáticas.

Entrando ya en los artículos recopilados en la obra, éstos están agrupados en cuatro secciones. La primera continúa estando dedicada a la cuestión de la lógica y los fundamentos de la matemática, y ofrece las contribuciones de Goldfarb (sobre la solitaria pugna de Poincaré contra los logicistas), de Friedmann (con un minucioso análisis de la repercusión de los teoremas de Gödel en la obra de Carnap) y de Moore, quien investiga la aparición del paradigma de la lógica de primer orden dentro del programa fregeano y de sus contemporáneos. Ya la lectura de esta primera sección da una idea clara del profundo cambio operado desde la publicación de la obra clásica de Putnam y Benacerraf sobre la filosofía de las matemáticas (1964) hasta la del libro aquí comentado.

La segunda sección muestra el modo actual de trabajar en este campo, basado en la selección de ejemplos históricos de teorías matemáticas en el contexto en el que se produjeron. Edwards, recuperando la figura de Kronecker y reinterpretando su polémica con Cantor, Birkhoff y Bennett, reevaluando la influencia del Programa de Erlangen que planteó

Klein, así como Dauben, mostrando la importancia del "Analysis non Standard" de Robinson para la cuestión de la fundamentación de las matemáticas, son tres excelentes ejemplos de este nuevo modo de hacer.

El artículo de Askey supone una transición hacia la tercera sección, centrada en las relaciones entre la historia y la filosofía de las matemáticas. Aparte del incisivo artículo de Crowe (como todos los suyos), es de destacar el estudio de Lorraine Daston sobre las matemáticas mixtas del siglo XVIII, y en particular sobre la teoría de la probabilidad: rompiendo la separación entre matemáticas puras y matemáticas aplicadas, Daston muestra la interacción entre las aplicaciones de una teoría matemática y sus teorizaciones, tema éste que, no por obvio hoy en día, sobre todo a partir de la concepción estructural de Sneed y sus seguidores, deja de tener interés y de ser novedoso en el ámbito de las matemáticas. La aportación de Stein va en una dirección complementaria, mostrando la incidencia de las investigaciones filosóficas sobre la práctica de las matemáticas (Dirichlet, Riemann y Dedekind son los autores estudiados, así como Hilbert). Browder estudia las relaciones entre las matemáticas y otras ciencias, tanto naturales como computacionales.

Finalmente Kitcher expone sus conocidas tesis sobre el naturalismo matemático (véase su libro *The Nature of Mathematical Knowledge,* Oxford Univ. Press, 1983), enlanzando con la cuarta sección, dedicada al contexto social de la investigación matemática. Grabiner la inicia proponiendo una hipótesis general sobre el cambio científico, que ilustra tanto con debates actuales en torno a la inteligencia artificial como con ejemplos históricos extraídos del los siglos XVII, XVIII y XIX. Aspray cierra el libro con un excelente trabajo sobre el desarrollo de la matemática en los EEUU durante el primer tercio del siglo XX, insistiendo en la importancia de las decisiones políticas y sociales al respecto: el desarrollo del departamento de matemáticas de la Universidad de Princeton y la aparición del Instituto de Estudios Avanzados son las dos instituciones por él analizadas.

Dentro de la variedad de aportaciones y de perspectivas, la obra editada por Aspray y Kitcher supone un buen ejemplo de las nuevas tendencias en filosofía de las matemáticas, íntimamente ligadas a la renovación de la historia de las matemáticas. Aun echándose de menos un estudio detallado de otros grupos, como el francés de Bourbaki, este libro evidencia el agotamiento de todo un programa de investigación, que ha dominado durante casi todo el siglo, según el cual las matemáticas serían, en el fondo, una parte de la lógica. Si bien dicha línea de investigación tuvo escasa influencia en la comunidad de matemáticos, sí que incidió profundamente entre los lógicos y filósofos de las matemáticas. El cambio producido a partir de los años setenta está permitiendo encontrar en los últimos simposios y congresos internacionales sobre estas materias nuevos puntos de contacto entre filósofos, historiadores y matemáticos, rompiendo el relativo aislamiento en el que habían estado dichas comunidades durante buena parte del siglo XX. Se recupera así una vieja tradición, que parecía perdida en nuestra época: el interés que los grandes matemáticos mostraron siempre, tanto por la obra de sus predecesores como por la reflexión filosófica sobre las matemáticas. (J. E.)

POPULATION DYNAMICS IN VARIABLE ENVIRONMENTS, por Shripad Tuljapurkar. Springer Verlag; Berlín, 1990.

La demografía relaciona los hechos observables en los individuos con la dinámica de las poblaciones a lo largo del tiempo. La teoría clásica de Lotka (1907) y de Leslie (1945) constituye la piedra angular cuando la dinámica del proceso es lineal y las diferentes tasas demográficas se mantienen a lo largo del tiempo. Sin embargo, estas hipótesis que ayudan a simplificar los problemas y que, en su momento, permitieron un notable avance de la teoría, son poco realistas y, en la mayor parte de los problemas prácticos ya se trate de demografía humana o de dinámica de poblaciones de otras especies, las tasas distan mucho de permanecer constantes en el tiempo, lo que dificulta notablemente el tratamiento de los problemas.

En este libro, el autor ofrece una aproximación a la dinámica de poblaciones bajo condiciones en las que las tasas demográficas varían con arreglo a diferentes patrones, a veces imprevisibles. No se trata de un libro elemental, ni siquiera de un libro de lectura fácil; sin embargo, contiene enfoques que resultan muy sugerentes al especialista que se enfrenta con problemas similares y, sobre todo, a quienes se interesan en la modelización o simulación de procesos demográficos. Lástima que las 154 densas páginas dejan poco espacio para desarrollar algunos ejemplos que se apuntan en el capítulo 5.

Tras la introducción se dedica un breve espacio a revisar las teorías clásicas y los problemas de la variación temporal determinista, lo que permite al lector introducirse fácilmente en la notación que luego utiliza el autor, aunque ésta no siempre es coherente a lo largo del libro. Como indica en la introducción el propio autor: "... algunos símbolos significan cosas diferentes en distintas partes del libro".

En el capítulo de aproximación a las variaciones aleatorias de las tasas demográficas, se realiza un sugerente planteamiento de los tipos particulares de modelos utilizables. Algunos de estos tipos, analizados en el libro, son: los modelos aleatorios del tipo IID, en los que las tasas varían de manera totalmente aleatoria, condicionada por factores exógenos (tales como las fluctuaciones ambientales); los modelos de tipo ARMA, en los que las tasas varían con arreglo a un patrón en el que se puede esperar la existencia de autocorrelación; los modelos de tipo markoviano, en los que las tasas varían de acuerdo con una distribución de probabilidades que es invariante en el tiempo; los modelos semi-markovianos, que se distinguen de los anteriores en que los intervalos de paso entre estados no son constantes, sino que se ajustan, a su vez, a una distribución de probabilidad que depende de los estados inicial y final; los modelos basados en la teoría de catástrofes, de particular aplicación en aquellas condiciones en las que se producen cambios drásticos en el ambiente. En este tipo de modelos subyace la idea de que la población puede absorber los cambios pequeños que tienen lugar en el ambiente pero resulta vulnerable a los cambios de mayor entidad.

El enunciado de algunos ejemplos tomados de la bibliografía, que se ajustan a otros tantos tipos de modelos y que cubren dominios tan diferentes como la demografía humana, grandes mamíferos (alces y ungulados), plantas bianuales, plantas perennes y cambios morfológicos en plantas, deja al lector con el deseo insatisfecho de ver alguno de estos ejemplos algo más que simplemente enunciado. Bien es verdad que, más adelante, se dedican los dos últimos capítulos al desarrollo más extenso de otros tantos ejemplos.

Se consagra un capítulo al problema clásico de la estructura de edades en las poblaciones, aunque nada se dice de otro problema no menos importante como el de las poblaciones estructuradas por el tamaño (por ejemplo, muchas poblaciones de plantas perennes) que está mereciendo un interés cada vez mayor.

Aspectos tan importantes desde el punto de vista práctico y teórico como el análisis de sensibilidad de las tasas de crecimiento al efecto del ruido sobre las mismas o sobre la estructura de las poblaciones encuentran un apartado especial en el libro.

En suma, una obra de muy difícil lectura para los no especialistas, que trata de condensar en unas pocas páginas los avances más recientes de la demografía y lo consigue con desigual fortuna, ya que frente a desarrollos de notable atractivo el lector interesado puede encontrar lagunas notables y que en definitiva evidencia que desde los tiempos de los pioneros como Lotka, Volterra o más tarde Leslie, la demografía ha cambiado las formas (con el uso de ordenadores y refinados modelos de cálculo), pero apenas ha avanzado en el terreno de la comprensión de los fenómenos que regulan la dinámica de las poblaciones. (C. G.)

BIOLOGICAL OCEANOGRAPHY. AN EARLY HISTORY, 1870-1960, por Eric L. Mills. Cornell University Press; Ithaca, 1989.

Difícilmente se encontrará otra ciencia tan ligada al lugar. A pesar de que la investigación actual exige la trabazón del científico y su medio –astrónomo y observatorio, físico teórico y acelerador, biólogo y laboratorio–, sólo en oceanografía adquiere especial relevancia el marco físico: Kiel, Plymouth, Woods Hole o la Estación de Nápoles. Pocas disciplinas, también, nacieron con un fin práctico tan claro: paliar el hambre. Bacalao, arenque, anchoveta o sardina son, desde el comienzo de la ciencia, objeto constante de estudio. Una memoria publicada en abril de 1907 por la Real Academia de Ciencias y Artes de Barcelona se titulaba justamente "Necesidad de la oceanografía para las industrias de pesca".

No madrugó, sin embargo, España en la ciencia oceanográfica. Odón de Buen, quien andando el tiempo habría de ocupar los cargos de máxima responsabilidad académica y administrativa en ese terreno, perdió una ocasión de oro en su viaje de circunnavegación a bordo de la fragata *Blanca* en 1885. Cuando hacía años que en esos mismos mares del norte de Europa comenzaban a fraguarse las ideas fundamentales de la ecología marina (plancton, sucesión, distribución, factores limitantes, producción, etcétera), él se entretiene en un coleccionismo antañón y proclamas cartageneras. La dinámica del plancton –la ecología marina– se descubrió en Kiel, se afianzó en Plymouth y se vertebró en Woods Hole y en la Scripps Institution. Hubo, por supuesto, interpolaciones, como la cuña escandinava, que aclaró el papel de las bacterias descomponedoras. (Pese a los intentos de Santander, Palma o Málaga, España no contaría hasta los años cuarenta con una voluntad seria de desarrollo oceanográfico –Francisco García del Cid– y estudio sistemático del plancton: el fitoplancton nerítico por Ramón Margalef, con la voluntariosa anticipación limnológica de Pardo en Valencia.)

En Kiel, y en la década de los setenta de la centuria pasada, la sistemática había dejado paso a la fisiología. Interesaba conocer los mecanismos físico-químicos subyacentes bajo los procesos vitales. Victor Hensen da un paso más al preocuparse por la fisiología de los organismos marinos como un todo. Es la definición de ecología que luego pasaría a los manuales: fisiología al aire libre. Pero no fue fácil levantar la disciplina, cuyos rasgos generales han pasado ya a los diccionarios: los productores primarios del plancton constituyen la fuente básica de energía sobre la que se asienta la vida en el agua (mar, lagos); el fitoplancton aprovecha la luz solar para convertir dióxido de carbono, agua y sales disueltas en compuestos orgánicos. Ese proceso muestra, a imagen de los ecosistemas terrestres, fluctuaciones estacionales. En buena parte de los mares, con la luz creciente de la primavera el fitoplancton se desarrolla hasta que agota las sales en solución en el agua. Debido a ese consumo y a la labor paralela de los animales que se han ido multiplicando ante esa oferta alimentaria, la eclosión primaveral del fitoplancton decae hasta que vuelve a levantarse en otoño, gracias, ahora, a la mezcla de las aguas promovida por la bajada de la temperatura y el aumento de los vientos, mezcla que supone la renovación de los nutrientes en las aguas superficiales. Sigue luego un período invernal, quiescente para el sistema de producción marina. El ciclo anual del plancton básase, pues, en la interacción entre luz, nutrientes, circulación, termoclina, capacidad fotosintetizadora o clorofila presente y consumo animal.

El estudio de la biología del plancton se desarrolló casi de espaldas al movimiento darwinista. A Victor Hensen y sus colaboradores Karl Brandt y Hans Lohmann les preocupaba, en un comienzo, medir, cuantificar, empezando por la distribución, abundancia, alimentación, reproducción y distribución de los peces de interés comercial. Se propusieron estaciones permanentes y viajes oceanográficos. Hubo que idear los aparatos –redes y filtros– para recoger e izar las muestras y crear nuevos conceptos –columna de agua, determinación química del plancton–. Cayeron los primeros prejuicios: los mares eran pobres, sobre todo, los del Trópico. ¿A qué se debía?

Para responder a la pregunta urgía conocer los ciclos de los elementos en el mar. Era obvio que las plantas asimilaban el dióxido de carbono. Pero, ¿quién fijaba allí el nitrógeno? ¿Cómo volvía éste a la atmósfera? Poco a poco se fueron conociendo los procesos de nitrificación y desnitrificación, los de fijación, asimilación y amonificación. Se descubrió el papel del sedimento. Se experimentó con los distintos elementos para identificar cuál de los imprescindibles para la vida era el factor limitante último, si el nitrógeno, los fosfatos o los silicatos. Se midió también la función condicionadora de la luz y de la temperatura, las corrientes y el viento. Es decir, desde muy pronto, se buscó apuntalar el medio físico, donde se desarrollara la ecología descriptiva (diversidad y clasificación de las comunidades), la ecología trófica (producción primaria, producción secundaria e integración de los sistemas) y la ecología demográfica, o curvas de población en determinadas áreas y a lo largo del curso de las estaciones. Ideas generales que se forjarían a través del estudio de casos particulares, verbigracia, la sucesión del clímax de producción de diatomeas por el de *Ceratium* (un dinoflagelado).

Complementa este libro, y el autor lo reconoce, el de Robert McIntosh *The Background of Ecology*, ya reseñado por la revista. Ambos son imprescindibles para una aproximación seria a la historia de una disciplina joven que, con todo los ismos y exageraciones que se quiera, ha impregnado la cultura de nuestro tiempo. (L. A.)

THE PHILOSOPHY OF JOHN SCOTTUS ERIUGENA, por Dermot Moran. Cambridge University Press; Cambridge, 1989.

Sabemos muy poco de la ciencia carolingia. Contados son los que traspasan el tópico de unos monjes guardando la cultura del desmoronado imperio romano, a modo de islas en un piélago de barbarie. Ni siquiera parece inquietar nuestra curiosidad qué se enseñaba en las escuelas episcopales o en las abaciales. Desde Boecio (siglo VI) hasta la recuperación manuscrita de Aristóteles (siglo XII) la Cristiandad medieval resistió primero y propició después figuras poderosas: Alcuino, Hrabano Mauro, llamado "praeceptor Germaniae" (maestro de

Alemania), y, sobre todo, Juan Escoto Eriúgena (800-877).

No existe todavía en ese tiempo una separación tajante entre obras de filosofía, de teología y de ciencia. Un problema especialmente debatido es, justamente, el de las relaciones entre fe y razón, que se reconocían mutuamente necesarias, de acuerdo con la sentencia de resabio agustiniano "fides quaerens intellectum, intellectus quaerens fidem" (que en traducción libre venía a decir: se ha de ahondar en el contenido doctrinal del dogma y el norte último del conocimiento está en Dios). Había, pues, ilación, no mixtura de planos. El imperativo de la primera parte de la sentencia promueve la explicación del mundo, de la naturaleza, a través de los comentarios al relato bíblico de la creación.

Mientras Bagdad se ha convertido en el centro de radiación cultural islámica, en Occidente Carlos el Calvo eleva el tono de la corte con la incorporación del obispo Hincmaro y del gramático Lupus de Fernières. El monarca se interesa por la matemática y la teología, y lo vemos con frecuencia en disputas conciliares. Eriúgena, maestro de la escuela de Laon, famosa por su enseñanza del griego, pertenece también al círculo del monarca. En lo que aquí nos interesa, su principal obra es *Periphyseon,* escrito entre 860 y 866. Se trata de un diálogo prolijo entre el mentor y su alumno cuyo núcleo doctrinal se inscribe entre los comentarios al Génesis (el *Hexameron*, o glosa de los seis días de la creación del mundo).

Allí aborda también la división de las ciencias (general, la lógica, y tres ciencias particulares: práctica o ética, natural o lo que hoy llamamos ciencias naturales y teología) y la graduación de los seres. Para entender la división de éstos, conviene saber que, en Eriúgena, *natura* es lo mismo que *physis* para los presocráticos: el ser y el no ser, toda la realidad en sus múltiples formas. Concepto que, en los presocráticos y en Eriúgena, va unido al de verdad en el sentido de correr el velo que nos impide ver las cosas en su puridad.

Hay otros puntos que merecen la atención del filósofo al viejo estilo (sus contribuciones a la dialéctica, al idealismo o al neoplatonismo) y que son analizados por Moran. Pero está por reconstruir la ciencia carolingia, en la que Eriúgena tuvo parte destacada en astronomía y en historia natural. Ciencia que no pudo apoyarse en autoridades griegas, sino que tuvo que tejerse con los mimbres rudos de la propia observación y creación en una Europa atormentada. (L. A.)

Apuntes

La voz más autorizada de la romanización, Estrabón, englobaba los vascones en el fondo común de los montañeses del norte peninsular. Claudio Sánchez Albornoz defendió la ''vasconización'' de los territorios que hoy ocupa la entidad administrativa de Euskadi por los genuinos hablantes de ese idioma (pueblos de Navarra, Alto Aragón y norte de Cataluña). La opinión que hoy goza de mayor aceptación la esbozó ya Antonio Tovar: el vasco es la lengua aborigen de la España septentrional: de Gerona a La Coruña.

Acertijo geológico: ¿qué es lo que tiene más de un millón de metros cúbicos de trozos de roca, recorre en línea recta hasta 100 kilómetros o más de 50 metros por segundo y puede remontar cerros? El *sturzstrom* contestó en 1932 Albert Heim. A lo que nadie ha respondido todavía con convincente seguridad es al por qué de esa avalancha pétrea y su desconcertante movilidad: no se ve todos los días un fenómeno de esos, ni es fácil identificarlos. En la Tierra no hay previsto hasta 1997 ningún episodio provocado con fines científicos. Mientras tanto habrá que seguir analizando las imágenes obtenidas de *sturzstroms* en otros planetas y trabajando con las simulaciones por ordenador.

Acertijo astronómico: si nos paseamos por el espacio interestelar, tomamos un puñado de polvo, lo congelamos dentro de una bola de hielo y porosa y lo lanzamos contra el Sol, ¿qué hemos creado? Daré una pista: los cristales de hielo se evaporan y el polvo queda libre. El respondón precipitado dirá que se trata de un cometa, resto primerísimo, según se cree, de enjambres de planetesimales a partir de los cuales se forjaron los planetas exteriores de nuestro sistema solar. Pero, ¿es cierta esa correspondencia entre polvo interestelar y polvo cometario? Sí lo parece en el caso del cometa Halley. No todos los astrofísicos, sin embargo, están dispuestos a admitir que sea esa una ley general.

Nadie diría que corren buenos tiempos para la sistemática o clasificación de los organismos. Nos referimos a los grandes grupos que, a estas alturas, cabría suponer plenamente asentados. La verdad es que sólo existe anuencia general en zoología y botánica; el mundo de los microorganismos constituye todavía un pozo obscuro. En 1959 R. H. Whittaker alumbró los ''cinco reinos'' (Animalia, Plantae, Fungi, Protista y Monera). Se le criticó haber usado distinta vara para medir el alcance de unos y otros y se propusieron tres ''dominios'': Bacteria, Archaea y Eucharya. Ernst Mayr, para quien la clasificación anterior cae en el mismo error que denuncia, acaba de proponer la suya, que consta de dos dominios (Procariotas y Eucariotas), varios subdominios cada uno y algunos reinos.

Aunque descubierto en 1846 por William Lassell, se sabía muy poco de Tritón, satélite de Neptuno, hasta la aproximación al mismo, en agosto de 1989, del Voyager 2. La sonda descubrió lo que parecían erupciones enormes, tipo geyser, que alcanzan hasta los ocho kilómetros de altura, con un diámetro de columna de centenares de metros. Pero, ¿cómo puede un satélite frío, tan alejado del Sol, convertir el hielo en el gas eyectado por esos chorros? Más que geyseres serían, opinan otros, inmensas tormentas de polvo para cuya aplicación se invoca el mecanismo de efecto de invernadero.

Los aztecas habían desarrollado un finísimo sentido de la observación naturalista. Dieron así nombre a las distintas plumas que componen el ala de un ave: *hiceinitzti*, para las rémiges primarias; *hinitomitl*, para las filoplumas; *tapalcatl*, para las coberteras; *tlachcayotl*, para las secundarias, y *hiniquanhiotl* para las plúmulas. Cada parte de la pluma tenía su denominación, como tenían la suya las distintas formas de pico de las aves.

La falta de espacio, la acumulación incesante de nuevos pliegos, la contratación de personal encargado, la periódica desinfección, etcétera, apoyan la postura de quienes abogan por quemar los herbarios de todo el mundo. Semejante afán pirómano ha provocado una dura reacción. ¿Quemar los herbarios de Jussieu, Banks, Jacquin, los expedicionarios españoles? ¿Prescindir de una innegable fuente de información farmacológica? ¿Eliminar una de las principales vías para estudiar, a través de los estomas de las hojas, la evolución del clima en los últimos siglos? La tea, piensan otros, debiera convertir en ceniza archivos, registros y demás pruebas de fe para los siglos venideros. La polémica sigue encendida.